

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:59:25 ; Search time 1.90456 Seconds
(without alignments)
633.203 Million cell updates/sec

Title: US-10-688-058-2

Perfect score: 587

Sequence: 1 QKIVETRIENFNFLYK.....TPSASLEVIKDYLSLKRL 119

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 70609 seqs, 10134256 residues

Total number of hits satisfying chosen parameters: 70609

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA_New.*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79.5	13.5	860	7	US-11-019-711-59
2	76.5	13.0	550	7	US-11-052-554A-200
3	72	12.3	971	7	US-11-052-554A-3
4	71.5	12.2	1889	7	US-11-102-476-46
5	71	12.1	1732	6	US-10-055-877-147
6	68.5	11.7	134	6	US-10-793-626-438
7	68.5	11.7	1237	7	US-11-052-554A-95
8	68	11.6	5024	6	US-10-793-626-2364
9	67.5	11.5	212	6	US-10-793-626-2350
10	67.5	11.5	213	6	US-10-793-626-1622
11	67.5	11.5	341	6	US-10-793-626-226
12	67.5	11.5	532	6	US-10-793-626-546
13	67.5	11.5	583	6	US-10-793-626-1358
14	67	11.4	617	6	US-10-982-545-2
15	67	11.4	968	7	US-11-000-463-281
16	67	11.4	989	6	US-10-821-234-975
17	67	11.4	3635	7	US-11-019-711-47
18	66.5	11.3	857	7	US-11-052-554A-218
19	66.5	11.3	1189	7	US-11-074-176-134
20	66	11.2	477	7	US-11-089-551A-34
21	66	11.2	654	7	US-11-120-308-52
22	65.5	11.2	492	7	US-11-152-697-2
23	65.5	11.2	539	7	US-11-152-697-5
24	65.5	11.2	1424	7	US-11-102-476-4
25	65.5	11.2	1562	7	US-11-052-554A-211

26	65	11.1	1255	7	US-11-022-562-235	Sequence 235, App
27	65	11.1	1255	7	US-11-052-554A-265	Sequence 265, App
28	65	11.1	1255	7	US-11-052-554A-266	Sequence 266, App
29	65	11.1	1255	7	US-11-052-554A-267	Sequence 267, App
30	65	11.1	1255	7	US-11-052-554A-268	Sequence 268, App
31	65	11.1	1255	7	US-11-052-554A-269	Sequence 269, App
32	65	11.1	1255	7	US-11-052-554A-270	Sequence 270, App
33	65	11.1	1255	7	US-11-052-554A-271	Sequence 271, App
34	65	11.1	1255	7	US-11-052-554A-272	Sequence 272, App
35	65	11.1	1255	7	US-11-052-554A-273	Sequence 273, App
36	65	11.1	1255	7	US-11-052-554A-274	Sequence 274, App
37	65	11.1	1255	7	US-11-052-554A-275	Sequence 275, App
38	65	11.1	1279	6	US-10-957-880-3	Sequence 3, Appli
39	64.5	11.0	222	6	US-10-793-626-128	Sequence 128, App
40	64.5	11.0	593	7	US-11-120-308-54	Sequence 54, Appl
41	64.5	11.0	3194	7	US-11-052-554A-90	Sequence 90, Appl
42	64	10.9	495	6	US-10-821-234-1154	Sequence 1154, Ap
43	64	10.9	513	6	US-10-485-517-160	Sequence 160, App
44	64	10.9	1032	7	US-11-014-367-2	Sequence 2, Appli
45	64	10.9	1452	6	US-10-995-561-778	Sequence 778, App

ALIGNMENTS

RESULT 1

US-11-019-711-59

; Sequence 59, Application US/11019711

; Publication No. US20060009634A1

; GENERAL INFORMATION:

- ; APPLICANT: Kekuda, Ramesh
- ; APPLICANT: Alsobrook II, John P
- ; APPLICANT: Tchernev, Velizar T
- ; APPLICANT: Liu, Xiaohong
- ; APPLICANT: Svytek, Kimberly A
- ; APPLICANT: Patturajan, Meera
- ; APPLICANT: Grosse, William M
- ; APPLICANT: Lepley, Denise M
- ; APPLICANT: Burgess, Catherine E
- ; APPLICANT: Vernet, Corine A.M.
- ; APPLICANT: Li, Li
- ; APPLICANT: Gorman, Linda
- ; APPLICANT: Edinger, Shlomit R
- ; APPLICANT: Sciore, Paul
- ; APPLICANT: Ellerman, Karen
- ; APPLICANT: Malyankar, Uriel M
- ; APPLICANT: Rothenberg, Mark
- ; APPLICANT: Stone, David J
- ; APPLICANT: Boldog, Ferenc L
- ; APPLICANT: Guo, Xiaojia
- ; APPLICANT: Shenoy, Suresh G
- ; APPLICANT: Anderson, David W
- ; APPLICANT: Padigar, Muralidhara
- ; APPLICANT: Taupier Jr, Raymond J
- ; APPLICANT: Miller, Charles E
- ; APPLICANT: Eisen, Andrew J

; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-235

; CURRENT APPLICATION NUMBER: US/11/019,711

; CURRENT FILING DATE: 2004-12-21

; PRIOR APPLICATION NUMBER: US/10/037,417

; PRIOR FILING DATE: 2002-09-20

; PRIOR APPLICATION NUMBER: 60/260,018

; PRIOR FILING DATE: 2001-01-05

; PRIOR APPLICATION NUMBER: 60/260,360

; PRIOR FILING DATE: 2001-01-08

; PRIOR APPLICATION NUMBER: 60/272,411

; PRIOR FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: 60/272,817

; PRIOR FILING DATE: 2001-03-02

; PRIOR APPLICATION NUMBER: 60/291,186

; PRIOR FILING DATE: 2001-05-15

; PRIOR APPLICATION NUMBER: 60/303,231

```
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 971
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-3

Query Match      12.3%; Score 72; DB 7; Length 971;
Best Local Similarity 28.4%; Pred.No.19;
Matches 27; Conservative 16; Mismatches 38; Indels 14; Gaps 3;

QY    20 KDESIVQLQDALSSATTSLSALQTQSNDRGSGLL-----SSFRLRKONNNHSHKDIS 70
     |||:| | | | | | | | :||:| | | | | | | | :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
DB    397 KDEATQAQAAKSASATT---ASTKATEAGSATATAAQSKSTAESAATRAETAAKRAEDIA 453
     |||:| | | | | | | | :||:| | | | | | | | :|:|:|:|:|:|:|:~

QY    71 NLTNLDLSLQSE--LARLKSNLNNEGMYFYTATPSA 103
     |||:| | | | | | | | :||:| | | | | | | | :|:|:~
DB    454 SAVALEDASTTKKGIVQLSSATNSTSESIAATPKA 488
     |||:| | | | | | | | :||:~

RESULT 4
US-11-102-476-46
; Sequence 46, Application US/11102476
; Publication No. US20050271680A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Pu
; APPLICANT: Palaniappan, Raghavan U.M.
; APPLICANT: Cornell Research Foundation, Inc.
; TITLE OF INVENTION: Novel Immunogenic Proteins of Leptosira
; FILE REFERENCE: 1153_080U51
; CURRENT APPLICATION NUMBER: US/11/102,476
; CURRENT FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: PCT/US2003/32385
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/417721
; PRIOR FILING DATE: 2002-10-10
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 1889
; TYPE: PRT
; ORGANISM: Leptospira interrogans
US-11-102-476-46

Query Match      12.2%; Score 71.5; DB 7; Length 1889;
Best Local Similarity 24.3%; Pred.No.49;
Matches 34; Conservative 21; Mismatches 40; Indels 45; Gaps 6;

QY    18 FYKDESLVQLQDA-----LSSATTLSLA-----LTQSN 45
     |||:| | | | | | | | :||:| | | | | | | | :|:~
DB    533 FSSDSSIATISNAQKNQAGYAATGTTDIKATFGKVSPVSTLSVTAAKLVEIQITPA 592
     |||:~

QY    46 NDRGSGLLSSF----LRKONNSHSDI SNRLTNLDSLSQELARLKSNLNNEGMFYTATP 101
     :||:~
DB    593 ASKAKGLTERFKATGFTDNSEN---DITNQVTWNSS--NTDIAEIKNTSGSKGITNTLTP 648
     :||:~

QY    102 -----SASLEVIKYIDL SYLK 116
     |||:~
DB    649 GSSEISNALGSIKSKVILK 668
     |||:~

RESULT 5
US-10-055-877-147
```


Sequence 147, Application US/10055877
Publication No. US20050288241A1
GENERAL INFORMATION:
APPLICANT: DeCristofaro, Marc
APPLICANT: Padigaru, Muralidhara
APPLICANT: Miller, Charles
APPLICANT: Tchernev, Velizar
APPLICANT: Zhong, Mei
APPLICANT: Anderson, David
APPLICANT: Ballinger, Robert
APPLICANT: Gerlach, Valerie
APPLICANT: Spytek, Kimberley
APPLICANT: Ratelli, Luca
APPLICANT: Kekuda, Ramesh
APPLICANT: Guo, Xiaojia
APPLICANT: Zerhusen, Bryan
APPLICANT: Andrew, David
APPLICANT: Mezes, Peter
APPLICANT: Patturajan, Meera
APPLICANT: Burgess, Catherine
APPLICANT: Eisen, Andrew
APPLICANT: Wolenc, Adam
APPLICANT: Baumgartner, Jason
APPLICANT: Shinkets, Richard
APPLICANT: Gusev, Vladimir
APPLICANT: Vernet, Corine
APPLICANT: Taupier Jr., Raymond
APPLICANT: Pena, Carol
APPLICANT: Shenoy, Suresh
APPLICANT: Li, Li
APPLICANT: Casman, Stacie
APPLICANT: Boldog, Ference
TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
FILE REFERENCE: 21402-251
CURRENT APPLICATION NUMBER: US/10/055,877
CURRENT FILING DATE: 2002-01-22
PRIORITY APPLICATION NUMBER: 60/262,892
PRIORITY FILING DATE: 2001-01-19
PRIORITY APPLICATION NUMBER: 60/263,598
PRIORITY FILING DATE: 2001-01-23
PRIORITY APPLICATION NUMBER: 60/263,799
PRIORITY FILING DATE: 2001-01-24
PRIORITY APPLICATION NUMBER: 60/264,117
PRIORITY FILING DATE: 2001-01-25
PRIORITY APPLICATION NUMBER: 60/264,139
PRIORITY FILING DATE: 2001-01-25
PRIORITY APPLICATION NUMBER: 60/264,478
PRIORITY FILING DATE: 2001-01-26
PRIORITY APPLICATION NUMBER: 60/263,351
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: 60/272,870
PRIORITY FILING DATE: 2001-03-02
PRIORITY APPLICATION NUMBER: 60/275,990
PRIORITY FILING DATE: 2001-03-14
PRIORITY APPLICATION NUMBER: 60/275,927
PRIORITY FILING DATE: 2001-03-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 512
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 147
LENGTH: 1732
TYPE: PRT
ORGANISM: Rattus norvegicus
US-10-055-877-147

Query Match 12.1%; Score 71; DB 6; Length 1732;
Best Local Similarity 24.1%; Pred. No. 49;
Matches 33; Conservative 25; Mismatches 53; Indels 26; Gaps 5;
Qy 1 OKIYVEIERRIENHFLFKDESLVQLQDALSSATTSI-----SALTQSNDRGSGLLSS 55
Db 594 OKLARHVRDKKEEVLVQKAESEQLRRAERAKKELEVHTAETASAKDRK----- 647

Qy 56 FLRKQNNHNSKDI--LRTLNDLSL-----QELARLKSNNNEGMFYATPS 102
Db 648 -LRQ-SRHYSKQLENELEGLKQKQISYSPGICSIHQEQEITKLTDLKKSIFYEEIS 705
Qy 103 ASLEVIKYDLSYLRRL 119
Db 706 KREGIHASEIKNLKEL 722
RESULT 6
US-10-793-626-438
Sequence 438, Application US/10793626
Publication No. US20050255478A1
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/10/793,626
CURRENT FILING DATE: 2004-03-04
PRIORITY APPLICATION NUMBER: 60/164,258
PRIORITY FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 438
LENGTH: 134
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: amino acid sequence
US-10-793-626-438

Query Match 11.7%; Score 68.5; DB 6; Length 134;
Best Local Similarity 20.9%; Pred. No. 3.3;
Matches 19; Conservative 23; Mismatches 40; Indels 9; Gaps 2;
Qy 18 FYKDESLVQLQDALSSATTSLSALTQSNDRGSGLLSFLRKQNNHNSKDI--LRT--- 74
Db 35 FGQERQQLILEANMADKMTSASSITHTNT-----LNNQIKNVKTDLGLDNLRTIKYK 89
Qy 75 -LNDLSLQELARLKSNNNEGMFYATPSAS 104
Db 90 DANDQVNRHQLEGGGQNDGLYFIIPSDAS 120

RESULT 7
US-11-052-554A-95
Sequence 95, Application US/11052554A
Publication No. US2005028866A1
GENERAL INFORMATION:
APPLICANT: Sachdeva, et al.
TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
FILE REFERENCE: 30853/40359A
CURRENT APPLICATION NUMBER: US/11/052,554A
CURRENT FILING DATE: 2005-02-07
PRIORITY APPLICATION NUMBER: US 60/589,227
PRIORITY FILING DATE: 2004-07-20
PRIORITY APPLICATION NUMBER: IN 173/DEL/2004
PRIORITY FILING DATE: 2004-02-06
NUMBER OF SEQ ID NOS: 763
SOFTWARE: Patent In version 3.3
SEQ ID NO 95
LENGTH: 1237
TYPE: PRT
ORGANISM: Helicobacter pylori J99
US-11-052-554A-95

Query Match 11.7%; Score 68.5; DB 7; Length 1237;
Best Local Similarity 24.6%; Pred. No. 56;
Matches 33; Conservative 17; Mismatches 47; Indels 37; Gaps 4;
Qy 6 EIERRIENHFLFYKDESLVQLQDAL---SSATTSLSALTQSNDRGSGLL----- 53

```

Db      193 QIEHSITKTSTTYAQSLLSNLTDAVNASSNNTTYVSALVNALNTLGVGFPTTSTHV 252
Qy      54 -----SSFLRKQSNNHKSIDSLSOELARLKSNLNN----- 92
Db      253 LNPFGQVFPYNSILGSTSSNNQQYNNLTLLMTLOGLS--TNQNNPNCANQIQ 310
Qy      93 --EGMEFYATPSAS 104
Db      311 CLEQFIQNLTPLA 324

RESULT 8
US-10-793-626-2964
; Sequence 2964, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1993-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2964
; LENGTH: 5024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
; NAME/KEY: MOD RES
; LOCATION: (5024)
; OTHER INFORMATION: variable amino acid
US-10-793-626-2964

Query Match 11.6%; Score 68; DB 6; Length 5024;
Best Local Similarity 29.3%; Pred. No. 3.7e+02;
Matches 27; Conservative 17; Mismatches 44; Indels 4; Gaps 3

Qy      21 DESLVOLODALGSATTSL-SALTQSNDRGSGLLSSFLRKQSNNHKSIDSLSNRLTND-- 77
Db      4148 DNAMKDLRITLESNTSPNSVNYINADKNQIEFDALQOASATSSKTSENPAETIEVL 4200
Qy      78 SLSQELARLKSNLNNEGMYFYTATPSASLEVIK 109
Db      4208 GLSQAIYDTKNALNGEQLRAT-EKSKDKLIK 4238

RESULT 9
US-10-793-626-2950
; Sequence 2950, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1993-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2950
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence

```

```
; NAME/KEY: MOD_RES
; LOCATION: (341)
; OTHER INFORMATION: variable amino acid
US-10-793-626-226

Query Match      11.5%; Score 67.5; DB 6; Length 341;
Best Local Similarity 27.2%; Pred. No. 14;
Matches 31; Conservative 18; Mismatches 34; Indels 31; Gaps 7;

Qy 21 DESLVQDALSATTSLTSALTSQSNDRG--SGLSSFLRKQNS-----NNHS----- 66
Db 183 DOALLQRLDLMK---FQLEELTEASLKEGEVDQLESIDIKRIQNSEKLNALNNAHQVLTD 239
Qy 67 -----KQISN-LRTNDSLSQELARLKNLNNEGMFYTATPSASLEVIKYD 111
Db 240 ESAIPDRLYELSNYLQTIINDIVPEKFLVKEDID---QFYTMLEDKAEI--YD 288

RESULT 12
US-10-793-626-546
; Sequence 546, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 546
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-546

Query Match      11.5%; Score 67.5; DB 6; Length 532;
Best Local Similarity 27.2%; Pred. No. 24;
Matches 31; Conservative 18; Mismatches 34; Indels 31; Gaps 7;

Qy 21 DESLVQDALSATTSLTSALTSQSNDRG--SGLSSFLRKQNS-----NNHS----- 66
Db 157 DOALLQRLDLMK---FQLEELTEASLKEGEVDQLESIDIKRIQNSEKLNALNNAHQVLTD 213
Qy 67 -----KQISN-LRTNDSLSQELARLKNLNNEGMFYTATPSASLEVIKYD 111
Db 214 ESAIPDRLYELSNYLQTIINDIVPEKFLVKEDID---QFYTMLEDKAEI--YD 262

RESULT 13
US-10-793-626-1358
; Sequence 1358, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1358
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-1358

; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-1358

Query Match      11.5%; Score 67.5; DB 6; Length 583;
Best Local Similarity 25.2%; Pred. No. 27;
Matches 29; Conservative 21; Mismatches 38; Indels 27; Gaps 5;

Qy 5 VEIERR---IEN--HNFLFYKDESIVQDALSATTSLTSALTSQSNDRGSGLLSSFLRK 59
Db 189 VOIENRVQIENLWINDITFNDELALYKELESSTQDKMIEKFFQFNDYGCCKILKSF--E 246
Qy 60 QNSNNHKKDINSIR---TLNDSLSQELARLKNLNNEGMFYTATPSASLEVIKYD 111
Db 247 EAKNKITKELDNLNHYKVNVELSENTKCLKA-----EKIKFD 284

RESULT 14
US-10-982-545-2
; Sequence 2, Application US/10982545
; Publication No. US20050244890A1
; GENERAL INFORMATION:
; APPLICANT: Davies, Huw Alun
; APPLICANT: McGuire, James
; APPLICANT: Simonsen, Anja Hviid
; APPLICANT: Blennow, Kaj
; APPLICANT: Podust, Vladimir
; APPLICANT: CIPHERGEN Biosystems, Inc.
; TITLE OF INVENTION: Biomarkers for Alzheimer's Disease
; FILE REFERENCE: 016866-011550US
; CURRENT APPLICATION NUMBER: US/10/982,545
; CURRENT FILING DATE: 2004-11-06
; PRIOR APPLICATION NUMBER: US 60/518,360
; PRIOR FILING DATE: 2003-11-07
; PRIOR APPLICATION NUMBER: US 60/526,753
; PRIOR FILING DATE: 2003-12-02
; PRIOR APPLICATION NUMBER: US 60/546,423
; PRIOR FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: US 60/547,250
; PRIOR FILING DATE: 2004-02-23
; PRIOR APPLICATION NUMBER: US 60/558,896
; PRIOR FILING DATE: 2004-04-02
; PRIOR APPLICATION NUMBER: US 60/572,617
; PRIOR FILING DATE: 2004-05-18
; PRIOR APPLICATION NUMBER: US 60/586,503
; PRIOR FILING DATE: 2004-07-08
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Secretogranin II (Chromogranin C, EM66,
; OTHER INFORMATION: secretoneurin) precursor
; NAME/KEY: PEPTIDE
; LOCATION: (182)-(214)
; OTHER INFORMATION: biomarker peptide M3680.7
US-10-982-545-2

Query Match      11.4%; Score 67; DB 6; Length 617;
Best Local Similarity 25.7%; Pred. No. 32;
Matches 28; Conservative 22; Mismatches 35; Indels 24; Gaps 5;

Qy 1 QKIVETIETRIENHFLFYKD-----ESLVQDALSATTSLTSALTSQSN 46
Db 223 QKLYTDDEDDIYKANNIAYEDVVGEDWNPVEEKIES--QTQEEVRDSKENIQNEIND 280
Qy 47 D-RGSGLLSSFLRKQNSNNHKKDINSIRTLNDSLSQELARLKNLNNEG 94
Db 281 EMKRSQGLG--IQBEDLRFKSKD-----QLSDSDVSKVIAYLRLVNAAG 322
```

```
RESULT 15
US-11-000-463-281
; Sequence 281, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chonghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIP4CN
; CURRENT APPLICATION NUMBER: US/11/000,463
; CURRENT FILING DATE: 2004-11-29
; PRIOR APPLICATION NUMBER: 10/291,265
; PRIOR FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/922,279
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 281
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-000-463-281
```

```
Query Match      11.4%; Score 67; DB 7; Length 968;
Best Local Similarity 23.3%; Pred No. 57;
Matches 41; Conservative 25; Mismatches 50; Indels 60; Gaps 8;

QY 1 QKIYVEIERRIENHNFYKDESLV---QLQDALSSATTSLSALTQSNDDR-GSGLLSFF 56
   : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 445 EDVALEEEETIYENEGFRPYAEDTLSARGSGSDSPGSSSL-LTRKNPKSGSPKSSSL 503
   : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 57 LR-----KQNSNNHVKDISNLTIN-----DSLQELARLKNLN----- 91
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 504 LKLKAERNAQAEKMGKHSASFSSITITNTCCSSSSSSSSSLSKTSGDLKPRASDAGI 563
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 92 -----NEGMYFTATPSASLEV-----IKYDLSYLKREL 119
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 564 RCTPKVRAKKDADANAGL--TSCPRAKSVRPKPFNRAESQSQEMDITLRRQL 617
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
```

Search completed: January 24, 2006, 20:55:32
Job time : 2.90456 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:57:15 ; Search time 20.9149 Seconds
(without alignments)
2377.329 Million cell updates/sec

Title: US-10-688-058-2
Perfect score: 587
Sequence: 1 QKIYVEIERRIENHFLFYK.....TPSASLEVIKYDLSYLRRL 119

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	587	100.0	119	5	US-10-688-058-2
2	566	96.4	407	5	US-10-688-058-4
3	91	15.5	326	3	US-09-943-002-8
4	90.5	15.4	2910	5	US-10-732-923-3342
5	88	15.0	503	5	US-10-450-763-50636
6	88	15.0	726	5	US-10-483-680-2
7	88	15.0	2400	5	US-10-450-763-50637
8	86	14.7	733	4	US-10-424-599-167699
9	86	14.7	1242	4	US-10-437-963-126385
10	86	14.7	1310	4	US-10-437-963-126423
11	86	14.7	1511	4	US-10-437-963-126422
12	85.5	14.6	523	5	US-10-450-763-50635
13	84	14.3	209	4	US-10-424-599-279739
14	84	14.3	602	5	US-10-732-923-13440
15	84	14.3	1309	5	US-10-450-763-60725
16	84	14.3	1394	6	US-11-097-143-9972
17	83	14.1	650	4	US-10-425-114-49905
18	83	14.1	1091	4	US-10-369-493-6328
19	82.5	14.1	761	4	US-10-416-330-35
20	82	14.0	2415	5	US-10-450-763-50638
21	81.5	13.9	343	4	US-10-425-115-204637
22	81	13.8	1832	4	US-10-437-963-126381
23	81	13.8	4624	4	US-10-408-765A-2991
24	80.5	13.7	666	3	US-09-801-368-36
25	80	13.6	186	4	US-10-724-972A-6870
26	80	13.6	497	4	US-10-425-115-279869
27	79.5	13.5	336	3	US-09-976-782-33

28	79.5	13.5	336	4	US-10-231-913-262	Sequence 262, App
29	79.5	13.5	860	4	US-10-080-334-166	Sequence 166, App
30	79.5	13.5	860	4	US-10-072-012-838	Sequence 838, App
31	79.5	13.5	860	4	US-10-037-417-59	Sequence 59, Appl
32	79.5	13.5	860	5	US-10-494-343-866	Sequence 866, App
33	79	13.5	570	4	US-10-425-114-57234	Sequence 57234, A
34	79	13.5	4080	4	US-10-307-817-138	Sequence 138, App
35	78.5	13.4	350	4	US-10-425-114-72202	Sequence 72202, A
36	78.5	13.4	950	4	US-10-282-122A-47285	Sequence 47285, A
37	78.5	13.4	1043	6	US-11-097-143-41328	Sequence 41328, A
38	78.5	13.4	1669	4	US-10-425-115-291649	Sequence 291649, Appl
39	78.5	13.4	1819	4	US-10-161-051-97	Sequence 97, Appl
40	78	13.3	589	4	US-10-425-114-68204	Sequence 68204, A
41	78	13.3	754	4	US-10-425-115-298211	Sequence 298211, A
42	78	13.3	4498	4	US-10-712-124-68	Sequence 68, Appl
43	78	13.3	4498	6	US-11-097-143-2577	Sequence 2577, Ap
44	77.5	13.2	918	6	US-11-097-143-6468	Sequence 6468, Ap
45	77.5	13.2	1246	6	US-11-097-143-19635	Sequence 19635, A

ALIGNMENTS

RESULT 1
US-10-688-058-2
; Sequence 2, Application US/10688058
; Publication No. US20050058661A1
; GENERAL INFORMATION:
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: HALE, KATHERINE S.
; APPLICANT: JOHNSTON, STEPHEN A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF THE GENUS
; TITLE OF INVENTION: BORRELIA
; FILE REFERENCE: MCRO:00305
; CURRENT APPLICATION NUMBER: US/10/688,058
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-688-058-2

Query Match	100.0%	Score	587	DB	5	Length	119
Best Local Similarity	100.0%	Pred. No.	3.7e-53				
Matches	119	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
QY	1	QKIYVEIERRIENHFLFYKDESIVQLODASSALTQSNNDRSGLLSFLRKQ	60				
DB	1	QKIYVEIERRIENHFLFYKDESIVQLODASSALTQSNNDRSGLLSFLRKQ	60				
QY	61	NSNHSKDISNRLTNDLSLSQELARKLNKNEGFFYATPSSASLEVIKYDLSYLRRL	119				
DB	61	NSNHSKDISNRLTNDLSLSQELARKLNKNEGFFYATPSSASLEVIKYDLSYLRRL	119				

RESULT 2
US-10-688-058-4
; Sequence 4, Application US/10688058
; Publication No. US20050058661A1
; GENERAL INFORMATION:
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: HALE, KATHERINE S.
; APPLICANT: JOHNSTON, STEPHEN A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF THE GENUS
; TITLE OF INVENTION: BORRELIA
; FILE REFERENCE: MCRO:00305
; CURRENT APPLICATION NUMBER: US/10/688,058
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 141

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-688-058-4

Query Match          96.4%; Score 566; DB 5; Length 407;
Best Local Similarity 95.8%; Pred. No. 3e-50;
Matches 114; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 QKIYVEIERRIENHFLFYKDESILVOLQDALSSATTSLSALTQSNDRGSGLLSFLRKQ 60
Db 174 EKIVYIEIRRIENHFLFYKDESILVOLQDALSSATTSLSALTQSNDRGSGLLSFLRKQ 233
QY 61 NSNNHKSQDISNRLTNDLSQELARLKSNNLNNEGMFYTATPSASLEVIKYDLSYLKRRL 119
Db 234 NSNNHKSQDISNRLTNDLSQELARLKSNNLNNEGMFYTATPSASLEVIKYDLSYLKEAL 292

RESULT 3
US-09-943-002-8
; Sequence 8, Application US/09943002
; Patent No. US20020045734A1
; GENERAL INFORMATION:
; APPLICANT: Duncan, Roy
; TITLE OF INVENTION: NOVEL REOVIRUS-DERIVED PROTEINS AND USES THEREFOR
; FILE REFERENCE: 78973-1C
; CURRENT APPLICATION NUMBER: US/09/943,002
; CURRENT FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 326
; TYPE: PRT
; ORGANISM: avian reovirus strain 138
US-09-943-002-8

Query Match          15.5%; Score 91; DB 3; Length 326;
Best Local Similarity 29.5%; Pred. No. 0.64;
Matches 28; Conservative 23; Mismatches 24; Indels 20; Gaps 3;

QY 23 SLVOLQDALSSATTSLSALTQSNDRGSGLLSFLRKQNSNNHKSQDISNRLTNDLSQEL 82
Db 85 TITTLQDSVSTLSITVDLTNT-----SSVHSEALSRLTIVDGNSTT 127
QY 83 LARLKSNNLNNEGMFYTATPS---ASLE-VIKYDLSY 114
Db 128 IDNLKSDVSSNGLAITDLQSRVKLSLESVSSHGLSF 162

RESULT 4
US-10-732-923-3342
; Sequence 3342, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3342
; LENGTH: 2910
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-732-923-3342

Query Match          15.4%; Score 90.5; DB 5; Length 2910;
Best Local Similarity 26.3%; Pred. No. 1.3;
Matches 30; Conservative 24; Mismatches 47; Indels 13; Gaps 4;
```

```
QY 8 ERRIENHNFYKDESILVOLQDALSSATTSLSALTQSNND---RGSGLLSFLRKQNSN 63
Db 424 EKSISNYNF-----DKTLFLFYNVLLKVSVHISGLYNTFMDILNTYSSDALNNHINSNNN 479
QY 64 NNSKDISNRLTND-SLSQELARLKSNNLNNEGMFYTATPSASLEVIKYDLSYLK 116
Db 480 NNNNNNNNNNDYFISYELRKLESIN---IYQRAVNAALYLFYILHMLE 529

RESULT 5
US-10-450-763-50636
; Sequence 50636, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIEP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 50636
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (267)..(315)
; OTHER INFORMATION: Receptor tyrosine kinase class II proteins domain identified
; OTHER INFORMATION: by eMATRIX, accession number BL002398, p-value=1.000e-40, raw sc
; OTHER INFORMATION: of 25.15
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (217)..(490)
; OTHER INFORMATION: Eukaryotic protein kinase domain identified by PFam,
; OTHER INFORMATION: accession name pkinase, E-value=9.6e-96, PFam score of 331.5
US-10-450-763-50636

Query Match          15.0%; Score 88; DB 5; Length 503;
Best Local Similarity 29.3%; Pred. No. 2.4;
Matches 41; Conservative 19; Mismatches 46; Indels 34; Gaps 9;

QY 4 YVEIERRIENHNFYKDESILV-----QLQDALSSATTSLSALTQSN-----NDRGS 50
Db 54 YFEIEKR-----LSHSQERLVNRETQCQSURLEKLNQKALTEKKNKELEIAQDRNI 107
QY 51 GLLSSFLR-KQNSNNHKSQDISNRLTNDLSQELARLKSNNLN--NEGMFYTATPSASLEV 107
Db 108 AIQSOFTTKKEEAKEKDL--IRT-NERLSQELYLETVKRLNEKLESNTTKGELQ- 163
QY 108 IKYD-----LSYLKRRL 119
Db 164 LKLDLQASDVSVKYREKRL 183

RESULT 6
US-10-483-680-2
; Sequence 2, Application US/10483680
; Publication No. US20050074839A1
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: Survivin Interacting Protein TPRI
; FILE REFERENCE: TPRI8W5
; CURRENT APPLICATION NUMBER: US/10/483,680
; CURRENT FILING DATE: 2004-01-13
```

; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-483-680-2

Query Match 15.0%; Score 88; DB 5; Length 726;
Best Local Similarity 29.3%; Pred. No. 3.8;
Matches 41; Conservative 19; Mismatches 46; Indels 34; Gaps 9;
QY 4 YVEIERRIENHFLFYKDESIV-----QLQDALSSATTSLSALTQSN-----NDRGS 50
DB 54 YFEIEKR-----LSHSQRLVNETRECOSLRLEKLNQKALKTEKKNKELEIAQDRNI 107
QY 51 GLLSFLR-KQNSNNHSDINSLRTINDLSLQELARLKSINL--NEGMPFYATPSASLEV 107
DB 108 AIQSQFTTKKELEAEKRD--IRT-NERLSQELELYLTVKRLNEKLSKESNTTKGELQ- 163
QY 108 IKYD-----LSYLKRL 119
DB 164 LKDELQASDVSVKYREKRL 183

RESULT 7
US-10-450-763-50637
; Sequence 50637, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIF3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 50637
; LENGTH: 2400
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (351)..(362)
; OTHER INFORMATION: 8 kw CHO2 ALPHA ANTIGEN PARAMYOSIN domain identified by
; OTHER INFORMATION: eMATRIX, accession number DM00588A, p-value=5.950e-09, raw score
; OTHER INFORMATION: 10.87
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2400)
; OTHER INFORMATION: xaa = x or * as defined in Table 2
US-10-450-763-50637

Query Match 15.0%; Score 88; DB 5; Length 2400;
Best Local Similarity 29.3%; Pred. No. 19;
Matches 41; Conservative 19; Mismatches 46; Indels 34; Gaps 9;
QY 4 YVEIERRIENHFLFYKDESIV-----QLQDALSSATTSLSALTQSN-----NDRGS 50
DB 83 YFEIEKR-----LSHSQRLVNETRECOSLRLEKLNQKALKTEKKNKELEIAQDRNI 136
QY 51 GLLSFLR-KQNSNNHSDINSLRTINDLSLQELARLKSINL--NEGMPFYATPSASLEV 107
DB 137 AIQSQFTTKKELEAEKRD--IRT-NERLSQELELYLTVKRLNEKLSKESNTTKGELQ- 192
QY 108 IKYD-----LSYLKRL 119

DB 193 LKDELQASDVSVKYREKRL 212
RESULT 8
US-10-424-599-167699
; Sequence 167699, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 167699
; LENGTH: 733
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(733)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_122449C.1.pap
US-10-424-599-167699

Query Match 14.7%; Score 86; DB 4; Length 733;
Best Local Similarity 23.4%; Pred. No. 6.3;
Matches 40; Conservative 28; Mismatches 49; Indels 54; Gaps 6;
QY 3 IYVEIERRI-----ENHNF-----LIFYKDESIVLQD 29
DB 323 LVIELNSRVSSLTLENNFRSKCDVMEKEYNVLFKSAEKKALDSKVLREKDEELHQLKD 382
QY 30 ----ALSATTSS---LSALTOSNDRGSLLS-----SFLRKQ-----NSNNHSD 68
DB 383 QFELALGEAKSQIVADIQSQRDDLKALDNSSKVNHLKQELQVQTQENLAKSRNESAE 442
QY 69 ISNRLTNDLSLQELARLKSINLNEGMPFYATPSASLEVIVKYDLSYLKRL 119
DB 443 LENLLTSLNKLCKELELEVSKLSSELTEVNESLQRLNLDADKHEAEMLASL 493

RESULT 9
US-10-437-963-126385
; Sequence 126385, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 126385
; LENGTH: 1242
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_28937C.1.pap
US-10-437-963-126385

Db 54 YFEIEKR-----LSHSQERLVNRETCQNLRLLEKLNQKVLTEKNEKLETAQDRNLGIQSOF 107
QY 51 GLLSSFLR-KQNSNNHSDKISNRLTNDLSQELARLKSNNLN--NEGMYTATPSASLEVIKYD-- 111
Db 108 AFQSQFTRKELEAKRDL--IRT-NERLSQEVYLTEDVKRLNEKLKESNTTKGELQ-LKLDL 169
QY 110 YDLS 113
Db 159 YPLT 162

RESULT 13
US-10-424-599-279739
; Sequence 279739, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Zhou Yihua
; APPLICANT: Kovalic David K
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 279739
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_94626C.1.pap
US-10-424-599-279739

Query Match 14.3%; Score 84; DB 4; Length 209;
Best Local Similarity 26.6%; Pred. No. 1.9;
Matches 25; Conservative 16; Mismatches 33; Indels 20; Gaps 4;
QY 39 SALTQSNDRGSGLLSSFLRKQ-----SNNHSDK--ISNRLTNDLSQELARL 86
Db 118 SGFVSDNNHDEAIRAFKNEHIQCKMEGFSSNHSQDTGLSNDRTTDTSS--VVSM 175
QY 87 KSNLNEGMF-----YTATPSASLEVIKYDLSY 114
Db 176 GRNNNRALYEDLEGPSVAPLSDLECLQWDDY 209

RESULT 14
US-10-732-923-13440
; Sequence 13440, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 13440
; LENGTH: 602
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-732-923-13440

Query Match 14.3%; Score 84; DB 5; Length 602;
Best Local Similarity 26.9%; Pred. No. 7.8;
Matches 36; Conservative 24; Mismatches 52; Indels 22; Gaps 8;
QY 4 YVEIERRI--ENHNFIFYKDESLVQLQDALSSATTSALTQSN-----NDRGSGLLSSF 56

Db 54 YFEIEKR-----LSHSQERLVNRETCQNLRLLEKLNQKVLTEKNEKLETAQDRNLGIQSOF 113
QY 57 LR-KQNSNNHSDKISNRLTNDLSQELARLKSNNLN--NEGMYTATPSASLEVIKYD-- 111
Db 114 TRAKEELEAKRDL--IRT-NERLSQEVYLTEDVKRLNEKLKESNTTKGELQ-LKLDL 169
QY 112 -----LSYLKRL 119
Db 170 QASDVTVKYREKRL 183

RESULT 15
US-10-450-763-60725
; Sequence 60725, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 60725
; LENGTH: 1309
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1055)..(1079)
; OTHER INFORMATION: FLAGELLIN SIGNATURE domain identified by eMATRIX, accession
; OTHER INFORMATION: number PR0207C, p-value=7.158e-23, raw score of 14.45
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (973)..(1113)
; OTHER INFORMATION: Bacterial flagellin N-terminus domain identified by Pfam, accession
; OTHER INFORMATION: accession name Flagellin_N, E-value=4.4e-78, Pfam score of 272.8
US-10-450-763-60725

Query Match 14.3%; Score 84; DB 5; Length 1309;
Best Local Similarity 23.4%; Pred. No. 22;
Matches 39; Conservative 24; Mismatches 50; Indels 54; Gaps 7;
QY 1 QKIYVEIERRIENHNFIFYKDESLVQ-----LQDAL-----S 32
Db 874 QRYPEQIKR---THNFRFYFKQLAQQYPEAVPLLYHLQTLLEHPELQLMQMEANYSQKQ 930
QY 33 SATTSLSALTQSNDRGSGLLSSFL---RKQNSNNHSDKISNRLT-----NDSL 79
Db 931 SLTLKMSAKSEANIDRFCELTQSWLPMDKTEKDPNQSSALSSSIERLSSGLRINSKADDA 990
QY 80 SOELA-RLKSNLN-----NEGMYTATPSASLEVIKYDLSYLK 116
Db 991 GOSTIANRFTSNIKGLTQARNANDGISVAQTTEGALSINNLRQVR 1037

Search completed: January 24, 2006, 20:54:26
Job time : 21.9149 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:30:59 ; Search time 7.3361 Seconds
(without alignments)
1341.095 Million cell updates/sec

Title: US-10-688-058-2
Perfect score: 587
Sequence: 1 QXIVYERIRIENHFLYK.....TPSASLEVIKYDLSYLKRL 119

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5-COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6-COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H-COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCITUS-COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE-COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	15.0	2349	2	US-09-538-092-914
2	82	14.0	294	2	US-09-248-796A-25451
3	80.5	13.7	666	2	US-09-487-558B-36
4	80	13.6	186	2	US-09-134-001C-4584
5	79.5	13.5	860	2	US-10-037-417-59
6	79	13.5	1290	2	US-09-248-796A-20654
7	78.5	13.4	333	2	US-09-248-796A-19426
8	78.5	13.4	1627	2	US-09-540-236-3533
9	76.5	13.0	214	2	US-09-248-796A-20889
10	76.5	13.0	367	2	US-09-543-681A-8041
11	76.5	13.0	2807	2	US-09-543-681A-4980
12	75.5	12.9	500	1	US-08-260-582-77
13	75.5	12.9	500	4	PCT-US95-05471-77
14	75.5	12.9	569	2	US-09-248-796A-16697
15	75	12.8	775	1	US-07-603-133B-13
16	75	12.8	1427	2	US-09-538-092-1044
17	74.5	12.7	2037	2	US-09-248-796A-22095
18	74.5	12.7	237	2	US-09-248-796A-18794
19	74.5	12.7	480	2	US-09-248-796A-20303
20	74	12.6	515	2	US-09-270-767-45839
21	74	12.6	733	2	US-09-232-338-10
22	73.5	12.5	353	2	US-09-248-796A-15118
23	73.5	12.5	415	2	US-09-543-681A-8002
24	73.5	12.5	592	2	US-09-861-451A-72
25	73.5	12.5	845	2	US-09-248-796A-17962
26	73	12.4	440	2	US-09-762-194-2
27	73	12.4	440	2	US-09-762-194-4

28	72.5	12.4	339	2	US-09-583-110-3067	Sequence 3067, Ap
29	72.5	12.4	348	2	US-09-107-433-2966	Sequence 2966, Ap
30	72.5	12.4	435	2	US-09-762-194-8	Sequence 8, Appl
31	72.5	12.4	435	2	US-09-762-194-10	Sequence 10, Appl
32	72.5	12.4	563	2	US-09-134-001C-3172	Sequence 3172, Ap
33	72.5	12.4	668	1	US-08-468-036-5	Sequence 5, Appl
34	72.5	12.4	668	1	US-08-376-843-5	Sequence 5, Appl
35	72.5	12.4	823	2	US-09-538-092-19	Sequence 19, Appl
36	72.5	12.4	823	2	US-09-248-796A-16699	Sequence 16699, A
37	72.5	12.4	1041	2	US-09-999-833A-498	Sequence 498, App
38	72.5	12.4	1041	2	US-09-954-987B-184	Sequence 184, App
39	72.5	12.4	1041	2	US-09-954-987B-186	Sequence 186, App
40	72.5	12.4	1041	2	US-10-020-445A-498	Sequence 498, App
41	72.5	12.4	1059	2	US-09-954-987B-187	Sequence 187, App
42	72.5	12.4	1111	2	US-09-914-259-28	Sequence 28, Appl
43	72.5	12.4	1177	2	US-09-134-001C-5106	Sequence 5106, Ap
44	72	12.3	290	2	US-09-134-000C-4994	Sequence 4994, Ap
45	72	12.3	580	2	US-09-248-796A-14473	Sequence 14473, A

ALIGNMENTS

RESULT 1
US-09-538-092-914
; Sequence 914, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Ioic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqFormatter Version 0.9
; SEQ ID NO 914
; LENGTH: 2349
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P12270
US-09-538-092-914

Query Match 15.0%; Score 88; DB 2; Length 2349;
Best Local Similarity 29.3%; Pred. No. 3.1;
Matches 41; Conservative 19; Mismatches 46; Indels 34; Gaps 9;
Qy 4 YVETRIENHNFYKDESLV-----QLQDALSSATTSLTQSN-----NDRGS 50
Db 54 YFELEK-----LSHQERLVNTRCQSLKLEKLNOLKALTEKKNKELETAQDRNI 107
Qy 51 GLSSFLR-KONSNNHSDISNLTNLSQELARKLSNLTN--NEGMYTATPSASLEV 107
Db 108 AIQSOFTRTKEELEAKRDL--INT-NERLSQELYLETVKRLNKLKESNTTKGELQ- 163
Qy 108 IKYD-----LSYLKRL 119
Db 164 LKLELQASDVSVKYREKRL 183

RESULT 2
US-09-248-796A-25451
; Sequence 25451, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al

```
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 25451
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-25451

Query Match 14.0%; Score 82; DB 2; Length 294;
Best Local Similarity 28.4%; Pred. No. 0.77;
Matches 33; Conservative 23; Mismatches 36; Indels 24; Gaps 6;

QY 1 QKIYVEIRRRHNHFLFYKDESIVQLQDALSSATTSLSALTQSNDR-----G 49
DB 142 QTVQSIDSFDESSVLYVDLEN-----EENLSSNSLSIV--NODHTYLQNKIHRISIN 195

QY 50 SGLSSFLRKQNSNNHSDISNRLTLN---DSLQELARLK--SNLANE-GMFT 98
DB 196 SYNLSDFVNRNNSNNNDNNANVEKNKSFELDQLQAWNRLTGINNANNVGEFT 251

RESULT 3
US-09-487-558B-36
; Sequence 36, Application US/09487558B
; Patent No. 6949356
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. 6949356man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofia
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.130
; CURRENT APPLICATION NUMBER: US/09/487,558B
; CURRENT FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/487,558
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 446
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-487-558B-36

Query Match 13.7%; Score 80.5; DB 2; Length 666;
Best Local Similarity 27.0%; Pred. No. 3.5;
Matches 24; Conservative 19; Mismatches 37; Indels 9; Gaps 3;

QY 13 NHHFLFYKDES----LVQLQDALSSATTSLSA-----LTOS--NNDRGSGLLSFLRKQNSN 63
DB 281 NKYYKKNDESGXPITQAPALPSTLMYTANPPYITQSPDNTWATGNTHVNNNNNS 340

QY 64 NHSKDISNRLTNDLSQELARLKNNLN 92
DB 341 NNSSNSNNNNNNNNNNNNNNNNNNNN 369
```

RESULT 4

```
US-09-134-001C-4584
; Sequence 4584, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4584
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4584

Query Match 13.6%; Score 80; DB 2; Length 186;
Best Local Similarity 27.0%; Pred. No. 0.67;
Matches 30; Conservative 27; Mismatches 32; Indels 22; Gaps 6;

QY 1 QKIYVEIRRRHNHFLFYKDESIVQLQD--ALSATTSLALTQSNDRGSGLLSFL- 57
DB 19 KKIYTEIQRLLETHF---RDTIDQISENTGISKAT-----IYRWKDKSSIIMSAFIE 70

QY 58 RKQNSNNHSDISNRLTNDLSQELARLKNNLNNEGMYTATPSASLEVI 108
DB 71 OSQYIVIHNQD-----NLYDDLFOFLVKIRD-----IYTKLGSVAIEIL 110

RESULT 5
US-10-037-417-59
; Sequence 59, Application US/10037417
; Patent No. 6903201
; GENERAL INFORMATION:
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Alsobrook II, John P
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Patturajan, Meera
; APPLICANT: Grosse, William M
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Li, Li
; APPLICANT: Gorman, Linda
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Sciore, Paul
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shenoy, Sureeh G
; APPLICANT: Anderson, David W
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Miller, Charles E
; APPLICANT: Eisen, Andrew J
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-235
; CURRENT APPLICATION NUMBER: US/10/037,417
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/260,018
; PRIOR FILING DATE: 2001-01-05
```

```
; PRIOR APPLICATION NUMBER: 60/260,360
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/272,411
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/272,817
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/291,186
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/303,231
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/305,060
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/318,405
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/318,700
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 860
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Myosin Tail
; OTHER INFORMATION: Consensus Sequence
US-10-037-417-59

Query Match          13.5%; Score 79.5; DB 2; Length 860;
Best Local Similarity 29.0%; Pred. No. 6.5;
Matches 29; Conservative 13; Mismatches 49; Indels 9; Gaps 3;

Qy 26 QIQDALSSATTSLSALTQSNDRGSGLLSFLRKQNSN-----NHSKDISNRLTINDSL 79
Db 184 QLESQSLVKQLDELQRLNDLTS--QKSRLQSENLDLTRLQLEAEAAQVSNLSKLSQ 241
Qy 80 SQELARLKSNNEGMFTATPSASLEVIKYDLSVLKRL 119
Db 242 ESQEEAKRSLEESR--ERANLQALRQLEHDLDSLREQ 280

RESULT 6
US-09-248-796A-20654
; Sequence 20654, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20654
; LENGTH: 1290
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (56), (60), (62), (65), (66), (109)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unkno
US-09-248-796A-20654

Query Match          13.5%; Score 79; DB 2; Length 1290;
Best Local Similarity 21.6%; Pred. No. 13;
Matches 29; Conservative 25; Mismatches 46; Indels 34; Gaps 4;

Qy 20 KDESIVQIQDALSSATTS-----LSALTQSNDRGSGLLSFLRK-----QNSNNHSDI 69
Db 593 EDELTPLEDNLSGTSSPDEKSFMTKLKEARDIESDGTMRKSVLRLMKLKQSERSTEDV 652
```

```
Qy 70 SNLRTINDLSLQSE-----LARKLSNLANEGMFTYATP-----SASL 105
Db 653 HNMLVLVDSMLQHVTVNESRVIGTDAHSVLTNIQKLDRLSTEDMARRAFAESKMLSQL 712
Qy 106 EVIKYDLSYLKRL 119
Db 713 ELVKEEKELLEKEL 726
```

RESULT 7

```
US-09-248-796A-19426
; Sequence 19426, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19426
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-19426
```

```
Query Match          13.4%; Score 78.5; DB 2; Length 333;
Best Local Similarity 23.6%; Pred. No. 2.2;
Matches 32; Conservative 23; Mismatches 36; Indels 17; Gaps 6;
```

```
Qy 18 FYKDES---LVO-LQDAL--SSATTSLSALTQSNDRGSGLLSFLRKQNSNNHSDISN 71
Db 228 FWKDDSATFLVQKLQALTRSESDIVASTSTNNQ-----SSFARDQQAANAIQIK 281
```

```
Qy 72 LRTINDLSLQELARLKSNNEGMFTATPSASLEVIKYDLSYLKRL 119
Db 282 LKLIABEENSQSKLPGNMNP--MILSVHAIATPKSV---FAFLKQL 324
```

RESULT 8

```
US-09-540-236-3533
; Sequence 3533, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3533
; LENGTH: 1627
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-3533
```

```
Query Match          13.4%; Score 78.5; DB 2; Length 1627;
Best Local Similarity 27.0%; Pred. No. 20;
Matches 24; Conservative 18; Mismatches 38; Indels 9; Gaps 2;
```

```
Qy 31 LSSATTSLSALTQSNDRGSGLLSFLRKQNSNNHSDISNRLTINDLSLQELARLKSNL 90
Db 882 LSSALVIDNTLKQSNNAKTQEIAN--LTQTVSGHTSQVRELGVTTGDLSSQKYSQVQTOA 938
```

```
Qy 91 NNEGMYFYAT-----PSASLEVIKYDLS 113
```



```

Query Match      12.9%; Score 75.5; DB 2; Length 569;
Best Local Similarity 23.8%; Pred. No. 10;
Matches 30; Conservative 25; Mismatches 44; Indels 27; Gaps 4

QY      15 NLFVKDRLVQLQDQLSSATTSLSALTQSNDR-----GSGLL 53
Db      194 NSILLSETSLSSDDTS--DTSISSTTSDTGNINAGSLHTSTASIKELSIQKTGVTL 252

QY      54 SSFKRKQNNHNSKDINSLRTLNDLSQELARKSLNNNEGMYTATPSASLEVIKYDLS 113
Db      253 SSYLSTKLSS--TSDITIELITTELTITLTIE--DNEPNTFTTPSSHSEIFSSD 307

QY      114 YLKRL 119
Db      308 VLSKQV 313

RESULT 15
US-07-603-133B-13
; Sequence 13, Application US/07603133B
; Patent No. 5298244
; GENERAL INFORMATION:
; APPLICANT: Redmond, Mark J.
; APPLICANT: Ijaz, Mohammed K.
; APPLICANT: Parker, Michael D.
; TITLE OF INVENTION: ASSEMBLED VIRAL PARTICLES AND THEIR
; TITLE OF INVENTION: USE IN A VACCINE TO ROTAVIRAL DISEASE
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/603,133B
; FILING DATE: 19901025
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Robins, Roberta L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9313-0004.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-7250
; TELEFAX: (415) 327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 775 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-07-603-133B-13

Query Match      12.8%; Score 75; DB 1; Length 775;
Best Local Similarity 27.4%; Pred. No. 17;
Matches 26; Conservative 15; Mismatches 40; Indels 14; Gaps 3

QY      32 SSATTSALTQSNND-----RGSGLLSFLRKQNNHNSKDINSLRTLNDLSQELAR 85
Db      554 SKLATSISEMTHLSLDAASSASRSVSRISNTISNWTNVSDNSVNTSLSDISTQTST 613

QY      86 LKSNL-----NNEGMYTATPSASLEVIKYDLS 113
Db      614 LSKNLKEMITQTEGMSFDDIGAAVLKT-KIDMS 647

```

Search completed: January 24, 2006, 20:02:37
Job time : 8.3361 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:27:53 ; Search time 25.5705 Seconds
(without alignments)
3283.383 Million cell updates/sec

Title: US-10-688-058-2

Perfect score: 587

Sequence: 1 QKIYVEIERRIENHFLYK.....TPSASLEVIKYDLSYLRRL 119

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	566	96.4	407	Q9R3K2 BORBU	Q9R3k2 borrelia bu
2	566	96.4	407	Q9S0G7 BORBU	Q9S0g7 borrelia bu
3	560	95.4	407	Q9S0A2 BORBU	Q9S0a2 borrelia bu
4	553	94.2	407	Q5XIN9 BORGA	Q5xyn9 borrelia ga
5	531.5	90.5	408	Q9S0S7 BORBU	Q9S0s7 borrelia bu
6	529	90.1	407	Q5XZ87 BORGA	Q5xz87 borrelia ga
7	508.5	86.6	341	Q5XYV6 BORGA	Q5xyv6 borrelia ga
8	300	51.1	404	Q9KKC1 BORHE	Q9kcl1 borrelia he
9	264.5	45.1	413	Q50931 BORBU	Q50931 borrelia bu
10	243	41.4	223	Q6ASJ2 BORGA	Q6asj2 borrelia ga
11	105	17.9	140	Q5XZ76 BORGA	Q5xz76 borrelia ga
12	94	16.0	581	Q7RQ99 PLAYO	Q7rq99 plasmodium
13	93	15.8	525	Q81IE7 PLAF7	Q8ie7 plasmodium
14	92.5	15.8	1056	Q7V111 HELHP	Q7v111 helicobacte
15	92	15.7	506	Q5SE55 DICDI	Q5se55 dictyosteli
16	91.5	15.6	1090	Q54Y90 DICDI	Q54y90 dictyosteli
17	91	15.5	326	Q9J1B3 GREOV	Q9j1b3 avian reovi
18	91	15.5	715	Q54G03 DICDI	Q54g03 dictyosteli
19	90.5	15.4	407	Q9WZHO THEMA	Q9wzho thermotoga
20	90.5	15.4	621	Q54PZ7 DICDI	Q54pz7 dictyosteli
21	90.5	15.4	2910	Q81BY8 PLAF7	Q81by8 plasmodium
22	90	15.3	646	Q6MTK5 MYCMS	Q6mtk5 mycoplasma
23	89.5	15.2	936	Q801N8 XENLA	Q801n8 xenopus lae
24	89.5	15.2	954	Q55FSA DICDI	Q55fs4 dictyosteli
25	89.5	15.2	3130	Q9BK46 PLAF7	Q9bk46 plasmodium
26	89.5	15.2	3130	Q81DX6 PLAF7	Q81dx6 plasmodium
27	89.5	15.2	3254	Q7YWE9 PLAF7	Q7ywe9 plasmodium
28	89.5	15.2	3254	Q9BK45 PLAF7	Q9bk45 plasmodium
29	89	15.2	648	Q54G04 DICDI	Q54g04 dictyosteli
30	89	15.2	3080	Q7YWF0 PLAF7	Q7ywf0 plasmodium
31	89	15.2	3081	Q7YWF1 PLAF7	Q7ywf1 plasmodium

32 89 15.2 3203 2 Q7YWE6 PLAF7
33 89 15.2 3203 2 Q7YWE7 PLAF7
34 88.5 15.1 1141 2 Q55EG1 DICDI
35 88.5 15.1 3096 2 Q7YWF2 PLAF7
36 88.5 15.1 3256 2 Q7YWE8 PLAF7
37 88 15.0 503 2 Q15655 HUMAN
38 88 15.0 687 2 Q58F23 HUMAN
39 88 15.0 692 2 Q504U6 HUMAN
40 88 15.0 726 2 Q15624 HUMAN
41 88 15.0 1681 2 Q81J39 PLAF7
42 88 15.0 2328 2 Q4PGJ7 USTMA
43 88 15.0 2349 1 TPR HUMAN
44 88 15.0 2363 2 Q99968 HUMAN
45 88 15.0 2363 2 Q5SWY0 HUMAN

ALIGNMENTS

RESULT 1
Q9R3K2 BORBU
ID Q9R3K2 BORBU PRELIMINARY; PRT; 407 AA.
AC Q9R3K2
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE Hypothetical protein BBL01 (Hypothetical protein BBS01) (Hypothetical
protein BEM01) (Hypothetical protein BPO1).
GN OrderedLocNames=BBL01, BEM01, BPO1, BBS01;
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid cp32-8, Plasmid cp32-3, Plasmid cp32-6, and Plasmid cp32-1.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_taxID=139;
RN [1]
NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31; PLASMID=cp32-8, cp32-3, cp32-6, and cp32-1;
RX MEDLINE=20138354; PubMed=10672174;
RA Casjens S., Palmer N., van Vugt R., Huang W.M., Stevenson B., Rosa P.,
RA Lathigra R., Sutton G.G., Peterson J.D., Dodson R.J., Haft D.H.,
RA Hickey E.K., Gwinn M.L., White O., Fraser C.M.;
RT "A bacterial genome in flux: the twelve linear and nine circular
extrachromosomal DNAs in an infectious isolate of the Lyme disease
spirochete Borrelia burgdorferi.";
RL Mol. Microbiol. 35:490-516(2000).
[2]
NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31; PLASMID=cp32-8, cp32-3, cp32-6, and cp32-1;
RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson S.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Bowman C., Garland S.A., Fujii C., Cotton M.D.,
RA Artiach P., Bowman C., Smith H.O., Venter J.C.;
RA Roberts K.M., Hatch B., Utterback T.R., Watthey L., McDonald L.A.,
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
burgdorferi.";
RL Nature 390:580-586(1997).
DR EMBL; AE001580; AAF07646.1; -; Genomic DNA.
DR EMBL; AE001576; AAF07474.1; -; Genomic DNA.
DR EMBL; AE001578; AAF07561.1; -; Genomic DNA.
DR EMBL; AE001575; AAF07429.1; -; Genomic DNA.
DR TIGR; BBL01; -;
DR TIGR; BEM01; -;
DR TIGR; BPO1; -;
DR TIGR; BBS01; -;
DR InterPro; IPR009427; DUF1073.
DR Pfam; PF06381; DUF1073; 1.
KW Complete proteome; Hypothetical protein; Plasmid.
SEQUENCE 407 AA; 47037 MW; 6440A2466F5B9F23 CRC64;

Q7ywe6 plasmodium
Q7ywe7 plasmodium
Q55eg1 dictyosteli
Q7ywf2 plasmodium
Q7ywe8 plasmodium
Q15655 homo sapien
Q58f23 homo sapien
Q504u6 homo sapien
Q15624 homo sapien
Q81j39 plasmodium
Q4pgj7 ustilago ma
P12270 homo sapien
Q99968 homo sapien
Q5swy0 homo sapien

RESULT 3					
ID	Q9S0A2 BORBU PRELIMINARY;	PRT;	407 AA.		
AC	Q9S0A2;				
DT	01-MAY-2000 (TrEMBLrel. 13, Created)				
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)				
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)				
DE	Hypothetical protein BBO01.				
GN	OrderedLocusNames=BBO01;				
OS	Borrelia burgdorferi (Lyme disease spirochete) .				
OC	Plasmid cp32-7.				
OC	Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;				
OC	Borrelia burgdorferi group.				
OX	NCBI_TaxID=139;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RC	STRAIN=ATCC 35210 / B31;				
RC	MEDLINE=20138354; PubMed=10672174;				
RX	Casjens S., Palmer N., van Vugt R., Huang W.M., Stevenson B., Rosa P.,				
RA	Lathigra R., Sutton G.G., Peterson J.D., Dodson R.J., Haft D.H.,				
RA	Hickey E.K.; Gwinn M.L., White O., Fraser C.M.;				
RT	"A bacterial genome in flux: the twelve linear and nine circular				
RT	extrachromosomal DNAs in an infectious isolate of the Lyme disease				
RT	spirochete <i>Borrelia burgdorferi</i> ."				
RL	Mol. Microbiol. 35:490-516(2000).				
RN	[2]				
RP	NUCLEOTIDE SEQUENCE.				
RC	STRAIN=ATCC 35210 / B31;				
RC	MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;				
RX	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,				
RA	Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,				
RA	Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,				
RA	Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,				
RA	Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,				
RA	Gocayne J.D., Weidman J.F., Uitterback T.R., Watthey L., McDonald L.A.,				
RA	Artiach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,				
RA	Roberts K.M., Hatch B., Smith H.O., Venter J.C.;				
RT	"Genomic sequence of a Lyme disease spirochete, <i>Borrelia</i>				
RT	<i>burgdorferi</i> ."				
RL	Nature 390:580-586(1997).				
DR	EMBL; AE001579; AAP07602.1; -; Genomic_DNA.				
DR	TIGR; BBO01; -				
DR	InterPro; IPR009427; DUF1073.				
DR	Pfam; PF06381; DUF1073; 1.				
KW	Complete proteome; Hypothetical protein; Plasmid.				
SQ	SEQUENCE 407 AA; 47061 MW; 8E34A6BF392D15EE CRC64;				
Query Match		95.4%;	Score 560;	DB 2;	Length 407;
Best Local Similarity		94.1%;	Fred. No. 6.8e-39;		
Matches 112;		Conservative 4;	Mismatches 3;	Indels 0;	Gaps 0;
Qy	1 QKIVYETERRIENHFLFYKDSESLVQLDALSSATTLSALTQSNDRSGILSPFRKQ 60	:	:		
Dd	174 EKIVVEIERRIENHFLFYKDSESLVQLDALSSATTLSALTQGNDRSGILSPFRKQ 233	:	:		
Qy	61 NSNNHSDISNLRLTNDLSLSQELARKLNINNEGMYFTATPSASLEVIKYDL SylKRRLL 119	:	:		
Dd	234 NSNNHSDISNLRLTNDLSLAQELARKLNINNEGMYFTATPSASLEVIKYDL SylKEAL 292	:	:		

RESULT 4	
Q5XYN9 BORG	
ID Q5XYN9 BORG PRELIMINARY;	PRT; 407 AA.
AC Q5XYN9	
DT 25-OCT-2004 (TREMBLrel. 28, Created)	
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)	
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)	
DE Hypothetical protein.	
DE ORENAMES-BGP242;	
OS Borrelia garinii PBI.	
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;	
OC Borrelia burgdorferi group.	
OX NCBI TaxID=290434;	

```

RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PB1;
RA Gloeckner G., Lehmann R., Romualdi A., Pradella S.,
RA Schulte-Spechtel U., Schilhabel M., Wilske B., Suhnel J., Platzter M.;
RT "Comparative analysis of the Borrelia garinii genome.";
RL Nucleic Acids Res. 32:6038-6046(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RP STRAIN=PB1;
RA Gloeckner G., Schilhabel M., Lehmann R., Platzter M.;
RA Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY722930; AU86093.1; -; Genomic_DNA.
DR InterPro; IPR009427; DUF1073.
DR Pfam; PF06381; DUF1073; 1.
RW Hypothetical protein.
SQ SEQUENCE 407 AA; 46848 MW; 8C13E7AE0350AE79 CRC64;

Query Match 94.2%; Score 553; DB 2; Length 407;
Best Local Similarity 93.3%; Pred. No. 2.6e-38;
Matches 111; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 QKIYVEIERRIENHFLFYKDESLVQLQDALSSATTSLSALTQSNDRGSGILSSFLRKQ 60
DB 174 EKIVVEIERRIENHFLFYKDESLVQLQDALSSATTSLSALTQSNDRGSGILSSFLRKQ 233

QY 61 NSNNHSDISNRLTNDLSLQELARLKSNNLNNEGMFYTATPSASLEVIKYDLSYLKRL 119
DB 234 NSNNHSDISNRLSNDLSLAELARLKSNNLNNEGMFYTATPSASLEVIKYDLSYLKRL 292

RESULT 5
Q9S057 BORBU
ID Q9S057 BORBU PRELIMINARY; PRT; 408 AA.
AC Q9S057;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein BBN01.
GN OrderedLocNames=BBN01;
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=139;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=20138354; PubMed=10672174;
RA Castjens S., Palmer N., van Vugt R., Huang W.M., Stevenson B., Rosa P.,
RA Lathigra R., Sutton G.G., Peterson J.D., Dodson R.J., Haft D.H.,
RA Hickey E.K., Gwinn M.L., White O., Fraser C.M.;
RT "A bacterial genome in flux: the twelve linear and nine circular
RT extrachromosomal DNAs in an infectious isolate of the Lyme disease
RT spirochete Borrelia burgdorferi.";
RL Mol. Microbiol. 35:490-516(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Fraser C.M., Castjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Weidman J.F., Utterback T.R., Wattley L., McDonald L.A.,
RA Artach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
DR EMBL; AB001581; AAF07681.1; -; Genomic_DNA.
DR TIGR; BBN01; -.

DR InterPro; IPR009427; DUF1073.
DR Pfam; PF06381; DUF1073; 1.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 408 AA; 47038 MW; 5EB60167E175F272 CRC64;

Query Match 90.1%; Score 529; DB 2; Length 407;
Best Local Similarity 87.4%; Pred. No. 2.8e-36;
Matches 104; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

QY 1 QKIYVEIERRIENHFLFYKDESLVQLQDALSSATTSLSALTQSNDRGSGILSSFLRKQ 60
DB 174 EKIVVEIERRIENHFLFYKDESLVQLQDALSSATTSLSALTQSNDRGSGILSSFLRKQ 233

QY 61 NSNNHSDISNRLTNDLSLQELARLKSNNLNNEGMFYTATPSASLEVIKYDLSYLKRL 119
DB 234 NSNNHSDISNRLSNDLSLAELARLKSNNLNNEGMFYTATPSASLEVIKYDLSYLKRL 292

RESULT 7
Q5XYI6 BORG
ID Q5XYI6 BORG PRELIMINARY; PRT; 341 AA.
AC Q5XYI6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=BGPI45;
OS Borrelia garinii PB1.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
```

```

OC Borrelia burgdorferi group.
OX NCBI_TaxID=290434;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PBI;
RC Gloeckner G., Lehmann R., Romualdi A., Pradella S.,
RA Schulte-Spechtel U., Schilhabel M., Wilske B., Suhnel J., Platzner M.;
RT "Comparative analysis of the Borrelia garinii genome.";
RL Nucleic Acids Res. 32:6038-6046(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBI;
RA Gloeckner G., Schilhabel M., Lehmann R., Platzner M.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY729222; AAU85996.1; -; Genomic_DNA.
DR InterPro; IPR009427; DUF1073.
DR Pfam; PF06381; DUF1073; 1.
KW Hypothetical protein.
SQ SEQUENCE 341 AA; 39289 MW; EF138B65AD40F827 CRC64;

Query Match      86.6%; Score 508.5; DB 2; Length 341;
Best Local Similarity 85.0%; Pred. No. 1.2e-34;
Matches 102; Conservative 10; Mismatches 7; Indels 1; Gaps 1;

QY 1 QKIYVEIERRIENHNFYKDESLVQLQDALSSATTSLTQ-SNDRGSGLLSFLRK 59
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 176 EKIIYEIEKRIENHNFYKDESLVQLQDALSSATTSLTQNSNKGSGLLSFLRK 235
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 60 QNSNNHSDISNLTNLNDSLSQELARLKSNNLNNEGMYTATPSASLEVIKYDLSYLKRL 119
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 236 QNSNNHSDLSNLSLNDLSLAHELLKNNLNNEGMYTANPSASLEVIKYDLSYLKEVL 295

RESULT 8
Q9KKC1 BORHE PRELIMINARY; PRT; 404 AA.
AC Q9KKC1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Borrelia hermsii.
OG Plasmid cp32.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=140;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HS-1;
RX MEDLINE=20316012; PubMed=10858201;
DOI=10.1128/IAI.68.7.3900-3908.2000;
RA Stevenson B., Porcella S.F., Oie K.L., Fitzpatrick C.A., Raffel S.J.,
Lubke L., Schrumpf M.E., Schwan T.G.;
RT "The relapsing fever spirochete Borrelia hermsii contains multiple,
RT antigen-encoding circular plasmids that are homologous to the cp32
RL plasmids of Lyme disease spirochetes.";
RL Infect. Immun. 68:3900-3908(2000).
DR EMBL: AF123078; AAF28886.1; -; Genomic_DNA.
DR InterPro; IPR009427; DUF1073.
DR Pfam; PF06381; DUF1073; 1.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 404 AA; 46409 MW; 5567119DBB024671 CRC64;

Query Match      51.1%; Score 300; DB 2; Length 404;
Best Local Similarity 50.8%; Pred. No. 5.1e-17;
Matches 62; Conservative 24; Mismatches 28; Indels 8; Gaps 2;

QY 1 QKIYVEIERRIENHNFYKDESLVQLQDALSSATTSLTQNSNDRGS---GLLSPL 57
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 170 EKIIYEIEKRIENHNFYKDESLSHIKDALSTDSVTLEAKRSSDKNSPFKDFPTGFI 229
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 58 RQNSNNHSDISNLTNLNDSLSQELARLKSNNLNNEGMYTATPSASLEVIKYDLSYLKR 117
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 230 SKEDG-----DVSSLSRANDELMRELRLKSLDNDGIFYSCEPNASLEVIKYDLSYLKE 284

```

```

QY 118 RL 119
Db 285 AL 286

RESULT 9
O50931 BORBU PRELIMINARY; PRT; 413 AA.
ID O50931;
AC O50931;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein BBA38.
GN OrderedLocusNames=BBA38;
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp54.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OX NCBI_TaxID=139;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
Gwinn M.L., Dougherty B.A., Tomb J.-P., Fleischmann R.D.,
Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A.,
Artiach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
DR EMBL: AE000790; AAC66300.1; -; Genomic_DNA.
DR PIR; F70211; F70211.
DR TIGR; BBA38; -.
DR InterPro; IPR009427; DUF1073.
DR Pfam; PF06381; DUF1073; 1.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 413 AA; 47901 MW; 382654B2DF31241 CRC64;

Query Match      45.1%; Score 264.5; DB 2; Length 413;
Best Local Similarity 47.4%; Pred. No. 5.1e-14;
Matches 55; Conservative 26; Mismatches 28; Indels 7; Gaps 3;

QY 1 QKIYVEIERRIENHNFYKDESLVQLQDALSSATTSLTQNSNDRGSGLLSFLRKQ 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 186 EQIYLEIEKRIENHNFYKDEHLVGLVLEIAKEEINVLA--NSKGG-IFSTFFRAQ 241

QY 61 NSNNHSDISNLTNLNDSLSQELARLKSNNLNNEGMYTATPSASLEVIKYDLSYLK 116
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 242 EPN---KSFQALSSVSDLSRLSKIKNTLNNDGIFYTASENARLEVIKYDLEFLK 294

RESULT 10
Q6ASJ2 BORGA PRELIMINARY; PRT; 223 AA.
ID Q6ASJ2;
AC Q6ASJ2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BGA33;
OS Borrelia garinii.
OG Plasmid lp54.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OX NCBI_TaxID=29519;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBI;

```

```

RA Gloeckner G., Lehmann R., Romualdi A., Pradella S.,
RA Schulte-Spechtel U., Wilske B., Suhnel J., Platzter M.;
RT "Comparative analysis of the Borrelia garinii genome.";
DR EMBL; CP000015; AAT93792.1; -; Genomic_DNA.
DR InterPro; IPR009427; DUF1073.
DR Pfam; PF06381; DUF1073; 1.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 223 AA; 25591 MW; E210E27A83C386D4 CRC64;

Query Match 41.4%; Score 243; DB 2; Length 223;
Best Local Similarity 49.1%; Pred. No. 1.6e-12;
Matches 54; Conservative 21; Mismatches 29; Indels 6; Gaps 3;

QY 7 IERRIENHFLFYKDESLVQLQDLSALTQSNDRGSLLSFLRKQSNHNS 66
DB 1 MEKIRNYNFLFYKDEHLVGLTLESLELAKEINVLV--NNKKG-IFSFPGQEPN--- 54

QY 67 KDISNLTNDLSDELARLKSNNNEGMFYTATPSASLEVTKYDLSVLK 116
DB 55 KSFSALSSASSKLSIELSKIKSLTNNDGIFYTASENASPEVIKYDLEFLK 104

RESULT 11
QXZ76_BORGA
ID Q5XZ76_BORGA PRELIMINARY; PRT; 140 AA.
AC Q5XZ76;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=BG056;
OS Borrelia garinii PBI.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=290434;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBI;
RA Gloeckner G., Lehmann R., Romualdi A., Pradella S.,
RA Schulte-Spechtel U., Schillabel M., Wilske B., Suhnel J., Platzter M.;
RT "Comparative analysis of the Borrelia garinii genome.";
RL Nucleic Acids Res. 32:6038-6046(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBI;
RA Gloeckner G., Schillabel M., Lehmann R., Platzter M.;
RA Submitted (SEP-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY72919; AAU85906.1; -; Genomic_DNA.
DR InterPro; IPR009427; DUF1073.
DR Pfam; PF06381; DUF1073; 1.
KW Hypothetical protein.
SQ SEQUENCE 140 AA; 15945 MW; 4B9DD0E998A523DF CRC64;

Query Match 17.9%; Score 105; DB 2; Length 140;
Best Local Similarity 88.0%; Pred. No. 0.38;
Matches 22; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 95 MFYTATPSASLEVTKYDLSYLKRL 119
DB 1 MFYTATPSASLEVTKYDLSYLKRL 25

RESULT 12
Q7RQ99_PLAYO
ID Q7RQ99_PLAYO PRELIMINARY; PRT; 581 AA.
AC Q7RQ99;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=PY01203;
OS Plasmodium yoelii yoelii.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=17XNL;
MEDLINE=22255706; PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Anguoli S.V., Suh B.B., Kooij T.W., Perteza M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M.,
RA Florens L., Yates J.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria parasite Plasmodium yoelii yoelii.";
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL01000316; EAA20493.1; -; Genomic_DNA.
DR InterPro; IPR007275; YTH.
DR Pfam; PF04146; YTH; 1.
DR PROSITE; PS50882; YTH; 1.
KW Hypothetical protein.
SQ SEQUENCE 581 AA; 68268 MW; B41B699C83A6B986 CRC64;

Query Match 16.0%; Score 94; DB 2; Length 581;
Best Local Similarity 28.7%; Pred. No. 17;
Matches 31; Conservative 18; Mismatches 41; Indels 18; Gaps 5;

QY 2 KIYVE---IERRIENHFLFYKDESLVQLQDLSALTQSNDRGSG 51
DB 213 KFYLDQILINKNYTH---LYKSQVLNMDKMLDALVLSGEKYEKNISITKERENN 269

QY 52 LLSSFLRKQSNHNSKDISNLTNDLSDELARLKSNNNEGMFYTA 99
DB 270 TLSDMF--TNSDNNYKDISNRLNLSGIN---IHDDKNLNRSGIINTA 312

RESULT 13
Q8IIE7_PLAF7
ID Q8IIE7_PLAF7 PRELIMINARY; PRT; 525 AA.
AC Q8IIE7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PF11_0227;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3D7;
MEDLINE=22255705; PubMed=12368864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berrihan M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallom S.J., Suh B., Peterson J., Anguoli S.,
RA Perteza M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McRadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.G.;
RT "Genome sequence of the human malaria parasite Plasmodium falciparum.";
RL Nature 419:498-511(2002).
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AE014839; AAN35811.1; -; Genomic_DNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.

```

Search completed: January 24, 2006, 19:56:39
Job time : 28.5705 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	264.5	45.1	413	2	F70211	hypothetical prote
2	90.5	15.4	407	2	F72343	hypothetical prote
3	88	15.0	503	4	S23741	hypothetical TPR/T
4	88	15.0	2094	2	S33124	tpz protein - huma
5	85.5	14.6	3394	2	T18501	hypothetical prote
6	85	14.5	348	2	I36839	C9L protein - vari
7	84	14.3	602	1	TVRTRR	protein kinase (EC
8	83	14.1	736	1	VPXRPC	outer layer protei
9	83	14.1	1091	2	T31407	hypothetical prote
10	82	14.0	2061	2	T13751	transcription fact
11	81	13.8	309	1	S34198	IGF Fc receptor II
12	81	13.8	561	2	B90414	hypothetical prote
13	81	13.8	1474	2	T18281	hypothetical prote
14	80.5	13.7	666	2	S50452	hypothetical prote
15	80.5	13.7	790	2	F75372	c-di-GMP phosphodi
16	80	13.6	321	2	A42507	F5L protein - vacc
17	80	13.6	322	2	E36213	F5L protein - vacc
18	80	13.6	1302	1	JC6009	surface-located me
19	79.5	13.5	2712	2	T05113	hypothetical prote
20	79	13.5	348	2	T28467	major envelope ant
21	79	13.5	348	2	C72154	E5L protein - vari
22	79	13.5	415	2	QJ1212	Mob protein - Bac
23	79	13.5	415	2	S05987	plasmid recombinat
24	79	13.5	1051	2	T18351	lmpI protein - Myc
25	79	13.5	1365	2	T30822	lmpI protein - Myc
26	79	13.5	2573	2	D71614	hypothetical prote
27	78.5	13.4	496	2	D97728	thermostable carbo
28	78.5	13.4	950	2	E70203	exonuclease sbcc (
29	78	13.3	1036	2	B69368	hypothetical prote

Db 1854 YIEIKERLKIETKINKQEKYIIIOQDNLI-LDNFNSTTT-----TNNNN----- 1901

Qy 54 SSFLRKQNSNNHSDISNLTNDLSQELARLKSNN 92

Db 1902 -----NDNNNDNNNDNNNDNYQQFIHSLKANLEN 1933

RESULT 6

I36839

C9L protein - variola virus (strain India-1967)

C;Species: variola virus

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004

C;Accession: I36839

R;Blinov, V.M.

submitted to GenBank, November 1992

A;Reference number: A36859

A;Accession: I36839

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-348 <BLI>

A;Cross-references: UNIPROT:P33865; UNIPARC:UPI0000167901; GB:X69198; NID:g456758; PIDN:

C;Superfamily: phospholipase D, chordopox type

Query Match 14.5%; Score 85; DB 2; Length 348;

Best Local Similarity 24.8%; Pred. No. 4.7;

Matches 27; Conservative 22; Mismatches 34; Indels 26; Gaps 3;

Qy 13 NHNLFYKDESILVQLQDALSSATTSLSALTSQNDNRGSLSSFLRKQNSNNHSDISNLT 72

Db 67 NHTFWYNNK-----VIALATEDRTSGYISFFIKRVNISLTCLINISL 109

Qy 73 RTLNDLSQELARLKSNNHGMFYTATPSASLEVIKYDLS-----YLKR 117

Db 110 RYEDSGSYKGVSHLK-----DGIIVTTTNNISVKANIIDLTGRVCYLTR 153

..

RESULT 7

TVTRKR

protein kinase (EC 2.7.1.37) raf - rat

N;Alternate names: kinase-related transforming protein raf; raf proto-oncogene protein-e

C;Species: Rattus norvegicus (Norway rat)

C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 31-Dec-2004

C;Accession: B26126

R;Ishikawa, F.; Takaku, F.; Nagao, M.; Sugimura, T.

Mol. Cell. Biol. 7, 1226-1232, 1987

A;Title: Rat c-raf oncogene activation by a rearrangement that produces a fused protein-

A;Reference number: A26126; MUID:87172791; PMID:3550433

A;Accession: B26126

A;Molecule type: mRNA

A;Residues: 1-602 <ISH>

A;Cross-references: UNIPARC:UPI000011P8CA; GB:M15428; NID:g206546; PIDN:AAA42002.1; PID:

C;Genetics:

A;Gene: raf

C;Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; serine/threonin

F;301-567/Domain: protein kinase homology <KIN>

F;309-317/Region: protein kinase ATP-binding motif

F;329/Active site: Lys #status predicted

F;453/Binding site: phosphate (Ser) (covalent) (by autophosphorylation) #status predicted

Query Match 14.3%; Score 84; DB 1; Length 602;

Best Local Similarity 26.9%; Pred. No. 11;

Matches 36; Conservative 24; Mismatches 52; Indels 22; Gaps 8;

Qy 4 YVEIERRI-ENHNFYKDESILVQLQDALSSATTSLSALTSQNN-----NDRGSLSSP 56

Db 54 YFEIEKLSQSOERLVNRETCQNLRLLEKLNQVKVLTENKKELETAQDENLQIOSQF 113

Qy 57 LR-QNSNNHSDISNLTNDLSQELARLKSNNL--NEGFMFYTATPSASLEVIKYD-- 111

Db 114 TRAKEEAEKRDLL--IRT-NERLSQEVYLTEDVKRLNEKLKESNTTKGELQ-LKIDEL 169

Qy 112 -----LSYLKERL 119

Db 170 QASDVTVKYREKRL 183

RESULT 8

VPXRPC

outer layer protein VP4 - porcine rotavirus C (strain Cowden)

N;Alternate names: hemagglutinin; outer capsid protein VP4

N;Contains: outer capsid protein VP5; outer capsid protein VP8

C;Species: porcine rotavirus C

C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004

C;Accession: D40822

R;Bremont, M.; Juste-Lesage, P.; Chabanne-Vautherot, D.; Charpilienne, A.; Cohen, J.

Virology 186, 684-692, 1992

A;Title: Sequences of the four larger proteins of a porcine group C rotavirus and compar

A;Reference number: A40822; MUID:92124743; PMID:1310192

A;Accession: D40822

A;Molecule type: genomic RNA

A;Residues: 1-736 <BRE>

A;Cross-references: UNIPROT:P26193; UNIPARC:UPI0000138C59; GB:M74218; NID:g333317; PIDN:

C;Genetics:

A;Map position: segment 4

C;Superfamily: rotavirus outer layer protein VP3

C;Keywords: Glycoprotein; hemagglutinin; outer capsid protein

F;1-243/Product: outer capsid protein VP8 #status predicted <VP8>

F;244-250/Region: cleavage processing #status predicted

F;251-736/Product: outer capsid protein VP5 #status predicted <VP5>

F;39,61,64,93,162,191,237,251,304,471,631/Binding site: carbohydrate (Asn) (covalent) #s

Query Match 14.1%; Score 83; DB 1; Length 736;

Best Local Similarity 22.2%; Pred. No. 18;

Matches 38; Conservative 29; Mismatches 46; Indels 58; Gaps 7;

Qy 3 IYVEIERRIENHNFYKDESILVQLQDALSSATTT-----SLSALTQSNDRGSL 53

Db 494 VRMDLERKINDLRNDFQLANSVALGILSLATSLTANLLESVPATASSVKDVAANVM 553

Qy 54 SSF-----LRKQNSNNH-----KDIN-----LRTLN----- 76

Db 554 KKFRTNFKRKATKAKYSEFIIGLLEDVTNVARNNGNFDITSAVMVSTTKQLTLD 613

Qy 77 -DLSLQELARLKSNN-----LNNEGFMFYTATPSASLEVIKYDLSYLKR 118

Db 614 VDTLSEIVARSADNFIPNRSYRMIEDGIVYEATPK---RTFSYDLTTLQQR 661

RESULT 9

T34107

hypothetical protein C18C4.5 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C;Accession: T34107

R;Gattung, S.

submitted to the EMBL Data Library, April 1996

A;Description: The sequence of C. elegans cosmid C18C4.

A;Reference number: Z21478

A;Accession: T34107

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1091 <GAT>

A;Cross-references: UNIPARC:UPI000017B7A1; ENML:U55369; PIDN:AAC25825.1; GSPDB:GN000023;

A;Experimental source: strain Bristol N2; clone C18C4

C;Genetics:

A;Gene: CESP:C18C4.5

A;Map position: 5

A;Introns: 38/3; 95/3; 179/2; 259/3; 301/3; 419/2; 573/3; 613/1; 875/3; 920/2; 959/2; 98

Query Match 14.1%; Score 83; DB 2; Length 1091;

Best Local Similarity 22.9%; Pred. No. 29;

Matches 36; Conservative 22; Mismatches 45; Indels 54; Gaps 6;

Qy 5 VVEIERRIENH-----FLFYKDESILVQLQDAL-----SSATTSLS 39

Db 722 IEMERILDNHNKLEKLEELKKSHTHTSLSVLREQNELAQDLQREEKQSSNLL 781
 QY 40 ALQTO--SNNDRGSLSSFLRKQNNH--SKDISNLR-----TLNDSLSQELARLKS 88
 Db 782 VLNQKIEKSEKERLEQIRSHTSQNSDTSKTIQDLEKISLELLKTDLLALDVQKLSK 841
 QY 89 NLNNEG-----FYTATPSPASLEV 107
 Db 842 SLDSKDQQLKEADEKNLMLEEVQALQNPATPSDAEI 878
 RESULT 10
 T13751
 transcriptions factor sin3 - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
 C:Accession: T13751
 R:Pennetta, G.; Pauli, D.
 Dev. Genes Evol. 208, 531-536, 1998
 A>Title: The Drosophila Sin3 gene encodes a widely distributed transcription factor esse
 A:Reference number: Z17753; MUID:99016223; PMID:9799435
 A:Accession: T13751
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-2061 <PEN>
 A:Cross-references: UNIPROT:O77025; UNIPARC:UPI000007DD75; EMBL:AJ007518; NID:g3281906;
 C:Genetics:
 A:Gene: sin3
 A:Cross-references: FlyBase:FBgn0022764
 C:Function:
 A:Description: is essential for embryonic viability

Query Match 14.0%; Score 82; DB 2; Length 2061;
 Best Local Similarity 24.4%; Pred. No. 78;
 Matches 29; Conservative 22; Mismatches 44; Indels 24; Gaps 4;
 QY 1 QKIYVEIERIENHNFYKDESLVQLQDALSSATTSLSALTQSNNDRGSLSSFLRKQ 60
 Db 1778 RQLFEHRRNRIRQHG-----TAASLESISVDSAISSNSTTGGGGSGSNNNN 1824
 QY 61 NNNHNS---KQISNLRTLNDSLSQELA-----RLKSNLNNEGMYFTATPSPASLEVIKYD 111
 Db 1825 NNNNSWGGKRLERLGAPOVLAQYAFNDRDEISTNISNGRCFVT---SKNLKLLKYD 1880

RESULT 11
 S34198
 IGE Fc receptor II, low-affinity - rat
 N:Alternate names: CD23; lymphocyte IGE receptor
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
 C:Accession: S34198
 R:Flores-Romo, L.; Shield, J.; Humbert, Y.; Graber, P.; Aubry, J.P.; Gauchat, J.F.; Ayal
 submitted to the EMBL Data Library, June 1993
 A:Description: Inhibition of an in vivo antigen-specific IGE response by antibodies to C
 A:Reference number: S34198
 A:Accession: S34198
 A:Molecule type: mRNA
 A:Residues: 1-309 <FLO>
 A:Cross-references: UNIPROT:Q63097; UNIPARC:UPI00000B624F; EMBL:X73579; NID:g313672; PID
 C:Superfamily: IGE receptor II; C-type lectin homology
 C:Keywords: B-cell; glycoprotein; immunoglobulin receptor; macrophage; tandem repeat; tr
 F:1-25/Domain: intracellular #status predicted <INT>
 F:14-22/Region: stop-transfer sequence
 F:24-46/Domain: transmembrane #status predicted <TM>
 F:47-309/Domain: extracellular #status predicted <EXT>
 F:126-309/Product: soluble IGE-binding factor (29K) #status predicted <ICI>
 F:149-309/Product: soluble IGE-binding factor (25-27K) #status predicted <BFI>
 F:164-283/Domain: C-type lectin homology <LCH>
 F:192-283,260-274/bisulfide bonds: #status predicted

Query Match 13.8%; Score 81; DB 1; Length 309;
 Best Local Similarity 27.6%; Pred. No. 9.1;

Matches 29; Conservative 20; Mismatches 40; Indels 16; Gaps 3;
 QY 21 DESLVQLQDA-----LSSATTSLSALTQSNNDRGSLSSFLRKQNNHNSKDISNLR 73
 Db 51 EKSLLQGLDAAIQNALQMSQNLBQLBQKQKQSDSQ-----SQNLNELQEDLINVK 104
 QY 74 TLNDSLSQELARLKS---NLNNEGMYFTATPSPASLEVIKYDLSYL 115
 Db 105 SQNSELSQNTLQEDLVNVKSGQLNEKRAASDSLEKLQEEVAKL 149

RESULT 12
 B90414
 hypothetical protein mcmA1 [imported] - Sulfolobus solfataricus
 C:Species: Sulfolobus solfataricus
 C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
 C:Accession: B90414
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F
 arett, R.A.; Ragan, M.A.; Seneen, C.W.; Van der Oost, J.
 submitted to GenBank, April 2001
 A:Description: Sulfolobus solfataricus complete genome.
 A:Reference number: A99139
 A:Accession: B90414
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-561 <KUR>
 A:Cross-references: UNIPROT:Q97W21; UNIPARC:UPI000006473C; GB:AE006641; NID:gl3815729; F
 C:Genetics:
 A:Gene: mcmA1
 C:Superfamily: Streptomyces isobutyryl-CoA mutase chain A

Query Match 13.8%; Score 81; DB 2; Length 561;
 Best Local Similarity 24.1%; Pred. No. 19;
 Matches 28; Conservative 21; Mismatches 51; Indels 16; Gaps 4;
 QY 4 YVEIERIENHNFYKDESLVQLQDALSSATTSLSALTQSNNDRGSLSSFLRKQNSN 63
 Db 133 YTEI--GVGVSMFPHWKEMDIVMSGIPLDKVTTSWTI-----NATAMELLSLIATSESR 185
 QY 64 NNSKDISNLRTLNDSLSQELARLKSNLNNEGMYFTATPSPA--SLEVIKYDLSYLKR 117
 Db 186 NINKSVLDGTVQNDILKEYIAR-----KNFIYPPSPSMRYAIDIIIESAKYLPK 234

RESULT 13
 T18281
 hypothetical protein D1 - slime mold (Dictyostelium discoideum)
 C:Species: Dictyostelium discoideum
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T18281
 R:Ribben Jr., W.K.; Gonzales, C.M.; Gonzales, S.T.; Pilkington, K.J.; Kiyosawa, H.; Hugh
 Genetics 148, 1117-1125, 1998
 A>Title: Dictyostelium discoideum nuclear plasmid Ddp5 is a chimera related to the Ddp1
 A:Reference number: Z14684; MUID:98198836; PMID:9539429
 A:Accession: T18281
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1474 <RIE>
 A:Cross-references: UNIPROT:Q23870; UNIPARC:UPI0000081644; EMBL:U00796; NID:g2702254; PI
 C:Genetics:
 A:Introns: 33/2

Query Match 13.8%; Score 81; DB 2; Length 1474;
 Best Local Similarity 23.1%; Pred. No. 63;
 Matches 36; Conservative 26; Mismatches 46; Indels 48; Gaps 7;
 QY 1 QKIYVEIERI-----ENHNFYKDESLV-----QL 27
 Db 911 KKLHLENKTAILLNEISSKQLYFDITISGYQYINYEYENCFSEKELQVSYAAGRVL 970
 QY 28 QDALSSATTSLSALTQSNNDRGSLSSFLRKQNNHNSKDISNLRTLNDSLSQELARLK 87

```
Db 971 KDKNDQINAELET--KNFKDFNLLNSL---KISNEHKSQDLNDLNTKNSLEKEIESLR 1025
QY 88 SNLNNEGMYTATPSASLEVIK--YDLSY----LKR 117
Db 1026 SRI-----IQLETTPTVSNQITQPAPEYSYKHEILKR 1057

RESULT 14
S50452
hypothetical protein YEL007w - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 09-Jul-2004
C:Accession: S50452
R:Dieckrich, F.S.
submitted to the EMBL Data Library, December 1994
A:Description: Saccharomyces cerevisiae chromosome V cosmid 9871, 8199, 9867, 9495 and
A:Reference number: S50428
A:Accession: S50452
A:Molecule type: DNA
A:Residues: 1-666 <DIE>
A:Cross-references: UNIPROT:P40002; UNIPARC:UPI000013AB94; EMBL:U18530; NID:G602367; PID
C:Genetics:
A:Gene: SGD:TOS9; MIPS:YEL007w
A:Cross-references: SGD:S0000733
A:Map position: 5L

Query Match 13.7%; Score 80.5; DB 2; Length 666;
Best Local Similarity 27.0%; Pred. No. 26;
Matches 24; Conservative 19; Mismatches 37; Indels 9; Gaps 3;

QY 13 NHNFLFYKDES----LVQLQDALSSATYLSA----LTQS--NDRGSLSSFLRKQNSN 63
Db 281 NKNYTYKNDSSGYITQFAPALPSTTLMTYANTPPYITQSPDNTNATGNTHTVNNNNNS 340

QY 64 NHSKDISNLRTLNDSLSQELARKSLNAN 92
Db 341 NNSNSNNNNNNNNNNNNNNNNNNNNNN 369

RESULT 15
F75372
c-di-GMP phosphodiesterase A - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: F75372
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vanathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: F75372
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-790 <WHI>
A:Cross-references: UNIPROT:Q9RTX3; UNIPARC:UPI00000D3E9E; GB:AE002007; GB:AE000513; NID
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1629
A:Map position: 1

Query Match 13.7%; Score 80.5; DB 2; Length 790;
Best Local Similarity 22.6%; Pred. No. 32;
Matches 28; Conservative 22; Mismatches 59; Indels 15; Gaps 2;

QY 4 YVEIERRIENHNFYKDESILVOLQDALSSATYLSALTQSNNDRGSLSSFLRKQNSN 63
Db 174 YAKLEQYDTAQPYLFEAHDHARQLGQFDMITSLNLACVQLSGGTGAALTLI--EEL 231

QY 64 NHSKDISNLRTLN-----DSLSQELARKSLNNEGMYTATPSASLEVIKY 110
Db 232 RHAKTVASPRYKSLVLFHFSQLLDLRLSEALTYCEQGLNTDSGFGTATPQCELQVTHG 291

Search completed: January 24, 2006, 19:58:58
Job time : 6.51452 secs
```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:21:03 ; Search time 26.5934 Seconds
(without alignments)
1966.133 Million cell updates/sec

Title: US-10-688-058-2

Perfect score: 587

Sequence: 1 QKIVVEIRRIENHFLYK.....TPSASLEVIKDYLSLKRL 119

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003s.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*
- 9: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	587	100.0	119	8	ADU98743
2	566	96.4	407	8	ADU98745
3	91	15.5	326	2	AY06109
4	88	15.0	503	4	ABG20277
5	88	15.0	726	6	ABP96033
6	88	15.0	2349	6	AAO16359
7	88	15.0	2400	4	ABG20278
8	85.5	14.6	523	4	ABG20276
9	84	14.3	1309	4	ABG30366
10	84	14.3	1984	4	ABE61060
11	83	14.1	650	8	ADN80539
12	83	14.1	1091	8	ADN23675
13	82.5	14.1	761	5	ABB77431
14	82.5	14.1	1396	8	ADN01130
15	82	14.0	2415	4	ABG20279
16	81	13.8	4624	7	ABE78572
17	81	13.8	4624	7	ADU71185
18	80	13.6	186	5	ABP39739
19	80	13.6	186	8	ADS07575
20	80	13.6	1141	7	ABM89610
21	79.5	13.5	336	5	ABM78804
22	79.5	13.5	356	6	ABM68241
23	79.5	13.5	860	5	ADI17302
24	79.5	13.5	860	7	ADC38517

25 79 13.5 570 8 ADX94570
26 79 13.5 2573 3 AAB18234
27 79 13.5 4080 8 ADO42289
28 78.5 13.4 350 8 ADY24418
29 78.5 13.4 950 6 ABU19361
30 78.5 13.4 1043 4 ABB71512
31 78.5 13.4 1627 8 ADL05847
32 78.5 13.4 1819 5 AAM47588
33 78 13.3 589 8 ADY12389
34 78 13.3 4498 4 ABB58595
35 77.5 13.2 388 3 AAB21232
36 77.5 13.2 918 4 ABB59892
37 77.5 13.2 1181 8 ADV89275
38 77.5 13.2 1181 8 ADV80528
39 77.5 13.2 1246 4 ABB64281
40 77.5 13.2 1250 3 AAY91279
41 77.5 13.2 1252 8 ADV82565
42 76.5 13.0 367 7 ADF07756
43 76.5 13.0 488 9 ADZ77216
44 76.5 13.0 550 9 AEB91490
45 76.5 13.0 565 4 AAB49642

ALIGNMENTS

RESULT 1
ADU98743
ID ADU98743 standard; protein; 119 AA.
XX
AC ADU98743;
XX
DT 24-FEB-2005 (first entry)
XX
DE Borrelia burgdorferi antigenic polypeptide seqid 2.
XX
KW antibacterial; vaccine; immune stimulation; immunity; antigen;
KW DNA library.
XX
OS Borrelia burgdorferi.
XX
PN WO2004103269-A2.
XX
PD 02-DEC-2004.
XX
PF 17-OCT-2003; 2003WO-US033056.
XX
PR 18-OCT-2002; 2002US-0419401P.
XX
PA (MACR-) MACROGENICS INC.
XX (TEXA) UNIV TEXAS SYSTEM.
XX
PI Sykes KF, Hale KS, Johnston SA;
XX
DR WPI; 2004-834155/82.
XX N-PSDB; ADU98742.
XX
PT Immunizing a subject against Borrelia burgdorferi infection comprises providing to the subject at least one Borrelia antigen or its fragment.
XX
PS Claim 27; SEQ ID NO 2; 121pp; English.
XX
CC The invention describes a method of immunizing a subject comprising providing to the subject at least one Borrelia antigen or its fragment to induce an immune response. Also described are: an isolated polynucleotide comprising a sequence having at least 17 contiguous nucleotides in common with a sequence not given in the specification; an isolated polypeptide having at least 5 consecutive amino acids of the sequence not given in the specification; a vaccine composition comprising at least one Borrelia antigen or at least one polynucleotide encoding a Borrelia antigen; screening for at least one test polypeptide or test polynucleotide encoding a polypeptide for an ability to produce an immune response; preparing a vaccine; vaccinating a subject; treating a subject infected

RESULT 6
AAO16359
ID AAO16359 standard; protein; 2349 AA.
XX
XX AAO16359;
DT 03-APR-2003 (first entry)
XX
DE Human translocated promoter region (TPR) protein, SEQ ID No 7.
XX
XX Human; p53 pathway modulating agent; p53-associated disorder;
KW translocated promoter region; TPR; cancer.
XX
OS Homo sapiens.
XX
XX WO200299050-A2.
XX
XX 12-DEC-2002.
XX
XX 03-JUN-2002; 2002WO-US017425.
XX
XX 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
XX
XX (EXEL-) EXELIXIS INC.
XX
XX Friedman L, Florman GD, Belvin M, Francis-Lang H;
PI WPI; 2003-167335/16.
XX
XX Identifying candidate p53 pathway-modulating agents, useful as
PT therapeutic targets for disorders associated with defective p53 function,
PT comprises screening for agents that modulate the activity of translocated
PT promoter region polypeptides.
XX
XX Claim 13; Page 84-94; 94pp; English.
XX
XX The invention comprises a method for identifying a candidate p53 pathway
CC modulating agent. The method involves providing an assay system
CC comprising a purified translocated promoter region (TPR) protein or
CC nucleic acid. The method of the invention is useful for identifying
CC candidate p53 pathway modulating agents for use as therapeutic targets
CC for disorders associated with defective p53 function (e.g. cancer). The
CC TPR proteins and nucleic acids are useful for identifying agents that
CC modulate TPR function. The present amino acid sequence represents a human
CC TPR protein
XX
XX Sequence 2349 AA;
SQ
Query Match 15.0%; Score 88; DB 6; Length 2349;
Best Local Similarity 29.3%; Pred. No. 24;
Matches 41; Conservative 19; Mismatches 46; Indels 34; Gaps 9;
QY 4 YVEIERRIENHFLFYKDESIV-----QLQDALSSATTSLTQSN-----NDRGS 50
DB 54 YFEIEKR-----LSHSQERLVNRETRECQSLRLEKLNQKALTEKRNKELEIAQDRNI 107
QY 51 GLLSSGFLR-KQNSNNHSDKISNRLTNDLSQELARLKSNLN--NEGMYFTATPSASLEV 107
DB 108 AIQSQFTTKKELEAEKEDL--IRT-NERLSQELELYLTEDVKRLNEKLESNTTGELQ- 163
QY 108 IKYD-----LSYLKRL 119
DB 164 LKLDLQASDVSVKYREKL 183
RESULT 7
ABG20278
ID ABG20278 standard; protein; 2400 AA.
XX

ABG20278;
AC
XX
XX 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #20269.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
PI WPI; 2001-639362/73.
DR N-PSDB; AAS84465.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 20; SEQ ID NO 50637; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have application in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: the sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 2400 AA;
SQ
Query Match 15.0%; Score 88; DB 4; Length 2400;
Best Local Similarity 29.3%; Pred. No. 24;
Matches 41; Conservative 19; Mismatches 46; Indels 34; Gaps 9;
QY 4 YVEIERRIENHFLFYKDESIV-----QLQDALSSATTSLTQSN-----NDRGS 50
DB 83 YFEIEKR-----LSHSQERLVNRETRECQSLRLEKLNQKALTEKRNKELEIAQDRNI 136
QY 51 GLLSSGFLR-KQNSNNHSDKISNRLTNDLSQELARLKSNLN--NEGMYFTATPSASLEV 107
DB 137 AIQSQFTTKKELEAEKEDL--IRT-NERLSQELELYLTEDVKRLNEKLESNTTGELQ- 192
QY 108 IKYD-----LSYLKRL 119
DB 193 LKLDLQASDVSVKYREKL 212


```
RESULT 8
ABG20276
ID ABG20276 standard; protein; 523 AA.
XX
AC ABG20276;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #20267.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR N-PSDB; AAS84463.
XX
New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 50635; 103pp; English.
XX
The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 523 AA;
Query Match 14.6%; Score 85.5; DB 4; Length 523;
Best Local Similarity 29.0%; Pred. No. 5.6;
Matches 36; Conservative 16; Mismatches 43; Indels 29; Gaps 7;
OY 4 YVEIERIENHFLFKDESLV-----QLQDALSSATTSLSALTQSN-----NDRGS 50
DB 54 YFEISR-----LSHQRLVNETRECOSLEKLNQKALTEKKNKELEIAQDRNI 107
OY 51 GLLSSFLR-KQNSNNHSDISNRLTINDLSQELARLXSNLNNEGMFYATPSPASLEVIK 109

RESULT 9
ABG30366
ID ABG30366 standard; protein; 1309 AA.
XX
AC ABG30366;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #30357.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR N-PSDB; AAS94553.
XX
New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 60725; 103pp; English.
XX
The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1309 AA;
Query Match 14.3%; Score 84; DB 4; Length 1309;
Best Local Similarity 23.4%; Pred. No. 28;
Matches 39; Conservative 24; Mismatches 50; Indels 54; Gaps 7;
```

```
QY 1 QKIYVEIERRIENHFLFYKDESLVQ-----LQDAL-----S 32
Db 874 QRYFQIKR---THNFRYFKQLAQOYPEAVPLYLHLQTLLEHPELQLM EANYSQKQK 930
QY 33 SATTSLSALTQSNDRGSLSSPL---RKQNNHSHKDINLRTL-----NDSL 79
Db 931 SLTKMSAKSEANIDRFCELTQSWLPMDKTEKDPNQSSALSSSIRLSGLRINS AKDDAA 990
QY 80 SOELA-RLKSNLN-----NEGMFYTATPSASLEVIKYDLSYLK 116
Db 991 GOSIANRFTTSIKGLTQARNANDGISVAQTTEGALS EINNLRQRV 1037

RESULT 10
ABB61060
ID ABB61060 standard; protein; 1984 AA.
XX
AC ABB61060;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 9972.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
DR N-PSDB; ABL05163.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 9972; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1984 AA;

Query Match 14.3%; Score 84; DB 4; Length 1984;
Best Local Similarity 30.0%; Pred. No. 50;
Matches 27; Conservative 16; Mismatches 39; Indels 8; Gaps 2;
QY 29 DALSSATTSLTQSNDRGSLSSFLRKQNNHSHKDINLRTLNDLSLQELARLKS 88
Db 1142 DANESSNDESATTSNCNTNQDEAAGKEQKQNNVSKDLRLKLSLEON----LARIQM 1197
QY 89 NLNNEGMFYTATPSASLEVIKYDLSYLKRR 118
```

```
Db 1198 MREN-----YDAGDBISELLKQWESLFLMQR 1223

RESULT 11
ADX80539
ID ADX80539 standard; protein; 650 AA.
XX
AC ADX80539;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polypeptide seqid 49905.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX
OS Unidentified.
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
PI WPI; 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
PS Claim 1; SEQ ID NO 49905; 15pp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 650 AA;
```

Query Match 14.1%; Score 83; DB 8; Length 650;
 Best Local Similarity 23.4%; Pred. No. 14;
 Matches 40; Conservative 28; Mismatches 49; Indels 54; Gaps 6;
 3 IYVEIERR-----IENHNF-----LFYKDESLVQLQD 29
 325 LVIELNSRVSSLTENNFRSCDVMEKYNDLKFSVAKKALDSKVLREKEEELHQLKD 384
 30 -----ALSAATTS---LSALTQNNDRGSGLLS-----SFLRKQ-----NSNNHSD 68
 385 QPEALGASKSQIVADLSQBRDDLKEALDNESKVNHLKQELQVTOENLAKSRNESAE 444
 69 ISNLTANDLSLQELARLKSNNLNEGMYTATPSASLEVIKVDLSYLKRL 119
 445 LENLLTLNKLCKELEVLKSLSSLTETVNESLQRLNLDKAEAEMLASEL 495

RESULT 12
 ADN23675
 ID ADN23675 standard; protein; 1091 AA.

ADN23675;
 02-DEC-2004 (first entry)
 Bacterial polypeptide #6328.
 Recombinant DNA construct; transformed plant; improved plant property;
 cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 pathogen tolerance; pest tolerance; plant disease resistance;
 cell cycle pathway modification; plant growth regulator;
 homologous recombination; seed oil yield; protein yield; carbohydrate;
 nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 bacterial polypeptide.
 Bacteria.

US2003233675-A1.
 18-DEC-2003.
 20-FEB-2003; 2003US-00369493.
 21-FEB-2002; 2002US-0360039P.
 (CAOY/) CAO Y.
 (HINK/) HINKLE G J.
 (SLAT/) SLATER S C.
 (CHEN/) CHEN X.
 (GOLD/) GOLDMAN B S.
 Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 WPI; 2004-061375/06.
 New recombinant DNA construct comprising a promoter positioned to provide
 for expression of a polynucleotide encoding a polypeptide from a
 microbial source, useful for producing plants with improved properties.
 Claim 1; SEQ ID NO 6328; 122pp; English.

The invention relates to a recombinant DNA construct comprising a
 promoter functional in a plant cell, where the promoter is positioned to
 provide for expression of a polynucleotide encoding a polypeptide from a
 microbial source. The invention also relates to a transformed plant
 comprising the recombinant DNA construct and a method of producing a
 transformed plant having an improved property. The plant is a crop plant
 such as maize or soybean. The method of producing a transformed plant
 having an improved property comprises transforming a plant with the
 recombinant DNA construct and growing the transformed plant, where the
 polynucleotide or polypeptide is useful for improving plant properties.
 The recombinant DNA construct is useful for producing plants with
 improved plant properties, e.g. improved cold, heat or drought tolerance,

CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 1091 AA;

Query Match 14.1%; Score 83; DB 8; Length 1091;
 Best Local Similarity 22.9%; Pred. No. 28;
 Matches 36; Conservative 22; Mismatches 45; Indels 54; Gaps 6;
 5 VEIERRIENHN-----FLFYKDESLVQLQD-----SSATTSLIS 39
 722 IEMERILDNNHNEKLEKREELKKSHTSHTSLESVLEEQNELAQIQDQLREEKEQSSNLL 781
 40 ALTQ--SNNDRGSGLLSFLRKQNNH--SKDTSNLR-----TLNDSLSQELARLKS 88
 782 VLNOKIEKSEKERLEQIRSHTSQNSDTSKTISDLEDKISELLKTNLDLALDVQLSK 841
 89 NINNEG-----FYTATPSASLEV 107
 842 SLDSKQOLKEADEKNLMLSEVQALQATPSDSAEI 878

RESULT 13

ABB77431
 ID ABB77431 standard; protein; 761 AA.

ABB77431;
 19-JUL-2002 (first entry)
 Human tumour marker protein se2-2.
 Human; tumour; cytostatic; cutaneous T cell lymphoma; CTCL; vaccine;
 antigen-presenting cell; tumour-specific T cell.
 Homo sapiens.
 WO200238803-A2.
 16-MAY-2002.

08-NOV-2001; 2001WO-DE004229.

08-NOV-2000; 2000DE-01055285.

(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

Eichmueller S, Schandendorf D, Usener D;

WPI; 2002-426959/45.

N-PSDB; ABL58958.

Composition containing tumor-associated nucleic acid, useful for
 diagnosis and treatment of tumors, especially cutaneous T cell lymphoma,
 also derived proteins and antibodies.

Claim 5; Fig 8; 84pp; German.

The invention relates to a diagnostic composition containing at least one
 of 23 nucleotide sequences (I, ABL58901-ABL58950) with altered expression
 associated with tumors. (I), including antisense sequences and
 ribozymes, also proteins (II, ABB77424-ABB77445) encoded by them and
 antibodies specific for (II), are useful for diagnosis, monitoring and

CC treatment of tumours, especially cutaneous T cell lymphoma (CTCL). (II)
 CC and antibodies to (II) are useful for vaccination. (II) can also be used
 CC to prepare pre-loaded antigen-presenting cells or tumour-specific T cells
 XX
 SQ Sequence 761 AA;
 Query Match 14.1%; Score 82.5; DB 5; Length 761;
 Best Local Similarity 24.3%; Pred. No. 19;
 Matches 28; Conservative 32; Mismatches 42; Indels 13; Gaps 4;
 QY 6 EIERIENHNFLFYKDESIVLOQDALSSATTSLSALTQSNDRGSGLLSSFLRKQNS-NN 64
 DB 161 ELKTQVEDLN-----ENLLKLEALTKSKNRENSLTDNLND----LNNELQKKQKAYNK 210
 QY 65 HSKDISNLTNDLSQELARLKSNNLNEGMYTATPSASLEVIKYDLSYLKRL 119
 DB 211 ILREKEEIDQENDELKQIKRLTSGL--QGKPLTDNKQSLIEELQKVKLENQL 263
 RESULT 14
 ADN01130
 ID ADN01130 standard; protein; 1396 AA.
 XX
 AC ADN01130;
 XX
 DT 01-JUL-2004 (first entry)
 XX
 DE Human cell growth, differentiation, and death-associated protein #14.
 XX
 KW human; cell growth; cell differentiation; cell death; CGDD;
 KW cell proliferative disorder; arteriosclerosis; atherosclerosis; bursitis;
 KW cirrhosis; hepatitis; polycythaemia vera; psoriasis; cancer;
 KW developmental disorder; Cushing's syndrome; hypothyroidism;
 KW neurological disorder; epilepsy; stroke; Alzheimer's disease;
 KW Pick's disease; Huntington's disease; Parkinson's disease;
 KW multiple sclerosis; autoimmune disorder; inflammatory disorder; AIDS;
 KW allergy; anaemia; asthma; contact dermatitis; diabetes mellitus;
 KW reproductive disorder; infertility; endometriosis; uterine fibroid.
 XX
 OS Homo sapiens.
 XX
 XX WO2004031364-A2.
 XX
 XX 15-APR-2004.
 XX
 XX 03-OCT-2003; 2003WO-US031441.
 XX
 XX 03-OCT-2002; 2002US-0416205P.
 XX
 XX 25-OCT-2002; 2002US-0421521P.
 XX
 XX 21-NOV-2002; 2002US-0428376P.
 XX
 XX 23-DEC-2002; 2002US-0436258P.
 XX
 XX 10-JAN-2003; 2003US-0439292P.
 XX
 XX 13-FEB-2003; 2003US-0447578P.
 XX
 XX (INCY-) INCYTE CORP.
 XX
 XX (BURR-) BURRILL J D.
 XX
 XX Elliott VS, Swarnakar A, Tang YT, Yue H, Lu DAM, Khare R,
 XX Chawla NK, Richardson TW, Marquis JF, Lal PG, Nguyen DB, Lee SY,
 XX Tran UK, Bhatia UG, Lee S, Blake JJ, Ho A, Zheng W, Gao J, Tran B,
 XX Yang YG, Gietzen KJ, Hafalia AJA;
 XX
 XX WPI; 2004-330172/30.
 XX
 XX N-PSDB; ADN01155.
 XX
 XX New isolated polypeptides associated with cell growth, differentiation
 XX and death, useful for diagnosing, treating or preventing e.g.
 XX PT atherosclerosis, psoriasis, cancers, Alzheimer's disease, AIDS, anemia,
 XX PT diabetes mellitus or infertility.
 XX
 XX Claim 1; SEQ ID NO 14; 213pp; English.
 PS
 XX The invention comprises the amino acid and coding sequences of human

CC proteins that are associated with cell growth, differentiation, and death
 CC (CGDD). The DNA and protein sequences of the invention are useful for
 CC diagnosing, treating or preventing disorders associated with aberrant
 CC expression of CGDD, such as: cell proliferative disorders (e.g.
 CC arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis,
 CC polycythaemia vera, psoriasis and cancers), developmental disorders (e.g.
 CC Cushing's syndrome and hypothyroidism), neurological disorders (e.g.
 CC epilepsy, stroke, Alzheimer's disease, Pick's disease, Huntington's
 CC disease, Parkinson's disease and multiple sclerosis),
 CC autoimmune/inflammatory disorders (e.g. AIDS, allergies, asthma,
 CC contact dermatitis and diabetes mellitus), and reproductive disorders
 CC (e.g. infertility, endometriosis and uterine fibroid). The present amino
 CC acid sequence represents a human CGDD-associated protein of the
 CC invention.
 XX
 SQ Sequence 1396 AA;
 Query Match 14.1%; Score 82.5; DB 8; Length 1396;
 Best Local Similarity 24.3%; Pred. No. 44;
 Matches 28; Conservative 32; Mismatches 42; Indels 13; Gaps 4;
 QY 6 EIERIENHNFLFYKDESIVLOQDALSSATTSLSALTQSNDRGSGLLSSFLRKQNS-NN 64
 DB 846 ELKTQVEDLN-----ENLLKLEALTKSKNRENSLTDNLND----LNNELQKKQKAYNK 895
 QY 65 HSKDISNLTNDLSQELARLKSNNLNEGMYTATPSASLEVIKYDLSYLKRL 119
 DB 896 ILREKEEIDQENDELKQIKRLTSGL--QGKPLTDNKQSLIEELQKVKLENQL 948
 RESULT 15
 ABG20279
 ID ABG20279 standard; protein; 2415 AA.
 XX
 AC ABG20279;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 XX Novel human diagnostic protein #20270.
 XX
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 XX Homo sapiens.
 XX
 XX WO200175067-A2.
 XX
 XX 11-OCT-2001.
 XX
 XX 30-MAR-2001; 2001WO-US008631.
 XX
 XX 31-MAR-2000; 2000US-00540217.
 XX
 XX 23-AUG-2000; 2000US-00649167.
 XX
 XX (HYSE-) HYSEQ INC.
 XX
 XX Drmanac RT, Liu C, Tang YT;
 XX
 XX WPI; 2001-639362/73.
 XX
 XX N-PSDB; AAS84466.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 XX PT diagnostics, forensics, gene mapping, identification of mutations
 XX PT responsible for genetic disorders or other traits and to assess
 XX PT biodiversity.
 XX
 XX Claim 20; SEQ ID NO 50638; 103pp; English.
 PS
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 2415 AA;

Query Match 14.0%; Score 82; DB 4; Length 2415;
Best Local Similarity 28.6%; Pred. No. 1.1e+02;
Matches 40; Conservative 19; Mismatches 47; Indels 34; Gaps 9;
QY 4 YVEIERRIENHFLPKDESIV-----QLQDALSSATTSLTQSN-----NDRGS 50
Db 116 YFEIEKR-----LSHSQRLVNETRECQSLRELEKLNQKALPERKKELEIAQDRNI 169
QY 51 GLLSSFLR-KQNSNNHSDISNLTINDSLSQELARLKSNLN--NEGMYFYATPSASLEV 107
Db 170 AIQSQFTRTKLEAEKRD--IRT-NERLSQELVLTEDVKRLNEKLESNTTKGELQ- 225
QY 108 IKYD-----LSYLKRRL 119
Db 226 LKLELQASDVSVKYRKRL 245

Search completed: January 24, 2006, 19:44:23
Job time : 29.5934 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:21:03 ; Search time 90.9538 Seconds
(without alignments)
1966.133 Million cell updates/sec

Title: US-10-688-058-4
Perfect score: 2062
Sequence: 1 MCDLRKTKLIDKISSLELYK.....QSSSFDNEELAILKEKLFSP 407

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003s.*
7: geneseqp2003bs.*
8: geneseqp2004s.*
9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2062	100.0	407	8	ADU98745
2	566	27.4	119	8	ADU98743 Borrelia
3	157.5	7.6	583	4	AG82132 S. epider
4	157.5	7.6	1009	6	ABU43217 Protein e
5	157.5	7.6	1010	5	ABP40333 Staphyloc
6	157.5	7.6	1010	8	AD804882 Staphyloc
7	153.5	7.4	448	6	ADA35569 Acinetoba
8	150.5	7.3	1957	8	ADN19417 Bacterial
9	143.5	7.0	1169	8	ADN18442 Bacterial
10	141	6.8	819	6	ABU19076 Protein e
11	140	6.8	703	6	ABU35605 Protein e
12	138	6.7	642	2	AAW80950 Amino aci
13	138	6.7	642	2	AAW62596 Human nuc
14	138	6.7	642	6	ABU03917 Human exp
15	138	6.7	642	6	ABU03918 Human exp
16	138	6.7	642	6	ABU03919 Human exp
17	138	6.7	642	6	ABU03921 Human exp
18	138	6.7	642	6	ABU03922 Human exp
19	138	6.7	642	7	ADN95337 Human BFC
20	138	6.7	642	8	ADL161248 Human pro
21	138	6.7	642	8	ADQ09256 Human HEC
22	138	6.7	642	8	ADRI4607 Human NF-
23	138	6.7	642	8	ABM81952 Tumour-as
24	138	6.7	642	9	ADX07596 Cyclin-de

25	138	6.7	642	9	ADY15272
26	138	6.7	664	7	ADC19752
27	137.5	6.7	1979	3	AAI18171
28	135.5	6.6	696	7	ADJ69261
29	133.5	6.5	2503	9	ADW44257
30	133.5	6.5	2543	9	ADW44255
31	133.5	6.5	2568	9	ADW44253
32	133.5	6.5	2633	4	ABG06505
33	133.5	6.5	2663	4	AAM39097
34	133.5	6.5	2663	8	ADQ17932
35	133.5	6.5	2663	9	ADX06867
36	133.5	6.5	2688	4	AAM40883
37	132.5	6.4	508	8	ADV69157
38	132.5	6.4	1769	8	ADX97546
39	132.5	6.4	1780	4	AAM38681
40	132.5	6.4	1780	4	ABU53201
41	131.5	6.4	878	5	ABP39533
42	131.5	6.4	878	8	ADS05472
43	131.5	6.4	1184	6	ABU25330
44	130.5	6.3	1073	4	AG67417
45	130.5	6.3	1073	6	ABG72692

ALIGNMENTS

RESULT 1
ADU98745
ID ADU98745 standard; protein; 407 AA.
XX
AC ADU98745;
XX
DT 24-FEB-2005 (first entry)
XX
DE Borrelia burgdorferi antigenic polypeptide seqid 4.
XX
KW antibacterial; vaccine; immune stimulation; immunity; antigen;
KW DNA library.
XX
OS Borrelia burgdorferi.
XX
PN W02004103269-A2.
XX
PD 02-DEC-2004.
XX
PF 17-OCT-2003; 2003WO-US033056.
XX
PR 18-OCT-2002; 2002US-0419401P.
XX
PA (MACR-) MACROGENICS INC.
(TEXA) UNIV TEXAS SYSTEM.
XX
PI Sykes KF, Hale KS, Johnston SA;
XX
DR WPI; 2004-834155/82.
XX
PT Immunizing a subject against Borrelia burgdorferi infection comprises providing to the subject at least one Borrelia antigen or its fragment.
XX
PS Claim 27; SEQ ID NO 4; 121pp; English.
XX
CC The invention describes a method of immunizing a subject comprising providing to the subject at least one Borrelia antigen or its fragment to induce an immune response. Also described are: an isolated polynucleotide comprising a sequence having at least 17 contiguous nucleotides in common with a sequence not given in the specification; an isolated polypeptide having at least 5 consecutive amino acids of the sequence not given in the specification; a vaccine composition comprising at least one Borrelia antigen or at least one polynucleotide encoding a Borrelia antigen; screening for at least one test polypeptide or test polynucleotide encoding a polypeptide for an ability to produce an immune response; preparing a vaccine; vaccinating a subject; treating a subject infected

CC with a pathogen; passive immunization; vaccination; detecting Herpesvirus
 CC ; and detecting antibodies to a herpesvirus. The method is useful in
 CC immunizing a subject against Borrelia burgdorferi infection. This is the
 CC amino acids sequence of a Borrelia burgdorferi polypeptide that may be
 CC useful in immunizing a subject against infection by a member of the
 CC Borrelia species.
 XX
 SQ Sequence 407 AA;

Query Match 100.0%; Score 2062; DB 8; Length 407;
 Best Local Similarity 100.0%; Pred. No. 1.8e-155;
 Matches 407; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCDLRKTLIDKISSLELYKYGIFFRNYIENVAEDCLNGLTLESAAHNVSEVELARLKV 60
 DB 1 MCDLRKTLIDKISSLELYKYGIFFRNYIENVAEDCLNGLTLESAAHNVSEVELARLKV 60
 QY 61 QLNALLNCIIISYRFGHGYVLVKTDLIDLEQPNVIELPIGFYLDYEVYRDIGVDFD 120
 DB 61 QLNALLNCIIISYRFGHGYVLVKTDLIDLEQPNVIELPIGFYLDYEVYRDIGVDFD 120
 QY 121 HITYKVKSNKNSLDVAKVHKSRLLIYENFDYILKRYVPCYTESFLDDIYLFKIIYVEI 180
 DB 121 HITYKVKSNKNSLDVAKVHKSRLLIYENFDYILKRYVPCYTESFLDDIYLFKIIYVEI 180
 QY 181 ERIENHNFLFYKDESLOQDALSSATTSLSALTQSNNDRGSGILSSFLRKQSNHNSK 240
 DB 181 ERIENHNFLFYKDESLOQDALSSATTSLSALTQSNNDRGSGILSSFLRKQSNHNSK 240
 QY 241 DISNLRNLNDSLSQELARLKSNNLNEGMYTATPSASLEVIKYDLSYLKEALALIKAKIG 300
 DB 241 DISNLRNLNDSLSQELARLKSNNLNEGMYTATPSASLEVIKYDLSYLKEALALIKAKIG 300
 QY 301 ADTKPLTRSFNEQAKLGNDGKDRSNYYDFLKGVEQVENACNLKTKYFGLDMKFN 360
 DB 301 ADTKPLTRSFNEQAKLGNDGKDRSNYYDFLKGVEQVENACNLKTKYFGLDMKFN 360
 QY 361 LIMLSEEQVERDIKLIELYSKYNOLIOSSFDNEELATLKEKLFSP 407
 DB 361 LIMLSEEQVERDIKLIELYSKYNOLIOSSFDNEELATLKEKLFSP 407

RESULT 2
 ADU98743
 ID ADU98743 standard; protein; 119 AA.
 AC ADU98743;
 XX
 XX
 XX
 XX 24-FEB-2005 (first entry)
 DE Borrelia burgdorferi antigenic polypeptide seqid 2.
 XX
 XX antibacterial; vaccine; immune stimulation; immunity; antigen;
 KW DNA library.
 XX
 XX Borrelia burgdorferi.
 XX
 XX WO2004103269-A2.
 PN
 XX
 XX 02-DEC-2004.
 XX
 XX 17-OCT-2003; 2003WO-US033056.
 PF
 XX
 XX 18-OCT-2002; 2002US-0419401P.
 PR
 XX
 XX (MACR-) MACROGENICS INC.
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX
 XX Sykes KF, Hale KS, Johnston SA;
 PI WPI; 2004-834155/82.
 XX
 XX DR N-PSDB; ADU98742.
 XX

PT Immunizing a subject against Borrelia burgdorferi infection comprises
 PT providing to the subject at least one Borrelia antigen or its fragment.
 XX
 XX Claim 27; SEQ ID NO 2; 121pp; English.

CC The invention describes a method of immunizing a subject comprising
 CC providing to the subject at least one Borrelia antigen or its fragment to
 CC induce an immune response. Also described are: an isolated polynucleotide
 CC comprising a sequence having at least 17 contiguous nucleotides in common
 CC with a sequence not given in the specification; an isolated polypeptide
 CC having at least 5 consecutive amino acids of the sequence not given in
 CC the specification; a vaccine composition comprising at least one Borrelia
 CC antigen or at least one polynucleotide encoding a Borrelia antigen;
 CC screening for at least one test polypeptide or test polynucleotide;
 CC encoding a polypeptide for an ability to produce an immune response;
 CC preparing a vaccine; vaccinating a subject; treating a subject infected
 CC with a pathogen; passive immunization; vaccination; detecting Herpesvirus
 CC ; and detecting antibodies to a herpesvirus. The method is useful in
 CC immunizing a subject against Borrelia burgdorferi infection. This is the
 CC amino acids sequence of a Borrelia burgdorferi polypeptide that may be
 CC useful in immunizing a subject against infection by a member of the
 CC Borrelia species.
 XX

SQ Sequence 119 AA;

Query Match 27.4%; Score 566; DB 8; Length 119;
 Best Local Similarity 95.8%; Pred. No. 4.2e-37;
 Matches 114; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 174 EKIVVEIERRIENHNFLFYKDESLOQDALSSATTSLSALTQSNNDRGSGILSSFLRKQ 233
 DB 1 QKIVVEIERRIENHNFLFYKDESLOQDALSSATTSLSALTQSNNDRGSGILSSFLRKQ 60
 QY 234 NSNNHSDKISNLRNLNDSLSQELARLKSNNLNEGMYTATPSASLEVIKYDLSYLKEAL 292
 DB 61 NSNNHSDKISNLRNLNDSLSQELARLKSNNLNEGMYTATPSASLEVIKYDLSYLKRRL 119

RESULT 3
 AAG82132
 ID AAG82132 standard; protein; 583 AA.
 XX
 XX AAG82132;
 XX
 XX
 XX 03-SEP-2001 (first entry)
 DT
 XX
 XX S. epidermidis open reading frame protein sequence SEQ ID NO:1358.
 DE
 XX Staphylococcus epidermidis SR1 strain; infection; diagnosis; vaccination;
 KW endocarditis.
 KW
 XX Staphylococcus epidermidis.
 OS
 XX WO200134809-A2.
 PN
 XX 17-MAY-2001.
 PD
 XX
 XX 09-NOV-2000; 2000WO-US030782.
 PF
 XX
 XX 09-NOV-1999; 99US-0164258P.
 PR
 XX (GLAX) GLAXO GROUP LTD.
 PA
 XX Kimmerly WJ;
 PI
 XX WPI; 2001-316495/33.
 DR N-PSDB; AAH52982.
 XX
 XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 PT useful for vaccinating against infections, e.g. endocarditis.
 XX
 XX Claim 18; Page 386; 2188pp; English.

CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AG81454 to AG83120, from *Staphylococcus epidermidis*. (I)
CC and (II) can have antibacterial activity and therefore can be used in
CC vaccination. The nucleic acids (I) may be used to produce the S.
CC epidermidis polypeptides (II) via the production of vectors containing
CC them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
CC represent oligonucleotide sequences and primers which are used in the
CC exemplification of the present invention. N.B. The present invention
CC specifically claims all the polynucleotide sequences given in the
CC sequence listing of the present specification, however the sequence
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
CC for SEQ ID NO:4455 to 4464
CC
XX Sequence 583 AA;

Query Match 7.6%; Score 157.5; DB 4; Length 583;
Best Local Similarity 25.6%; Pred. No. 0.0012;
Matches 80; Conservative 55; Mismatches 98; Indels 79; Gaps 16;
QY 131 KNSLDVAVKHKSLIYENFDYILKRVPCVTSFLLDIYLFKIVYEIER--IEN- 186
DB 154 QSNKDKQSILRT-LFNSERFDEI-----RHLLVENVKQEK--VQIENRYTOIENL 201
QY 187 -HNFLFYKDESIVQLQDALSSATTSLSALTSQNNDRGSGILSSFLRKQSNHNSKDINSL 245
DB 202 WNDIDTFNNDELALYKELESSQTDKMIKPPQFNDYGCKILKSF--EAKNKITKELDDL 259
QY 246 RN---LNDLSQELARLKSNNLNEGMYTATPSASLEVIKYD-----LSYLKEALAL 294
DB 260 NHKYNVVELSENTKQLKA-----EKIKFDLLKQEQYIDKLKQELKM 302
QY 295 I---KAKIGADTKEPLTRS FNEQAKGLGNDGKDRSNYYDFLKGVOEVENACN----- 345
DB 303 IQESKVLITYFRLQSLKDKDELVSLEHQSKLNETHYNEIKGFQKQLEHLSTRENEIT 362
QY 346 -----LKTQYGLDMKFNLSL--IMLSBEQK--VERDIKLIJELYSKYNOLI-----QS 389
DB 363 QFNQYLEKNQV-----FNQDKIISVQKQFVIEEEK--RLYSEYNDLITKKEELTKE 415
QY 390 SFDNEELAILK 401
DB 416 MNKNKDFAIIE 427

RESULT 4
ABU43217
ID ABU43217 standard; protein; 1009 AA.

XX AC ABU43217;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #28744.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS *Staphylococcus epidermidis*.
XX FN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR N-PSDB; ACA47087.
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
PS Claim 25; SEQ ID NO 71141; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1009 AA;

Query Match 7.6%; Score 157.5; DB 6; Length 1009;
Best Local Similarity 25.6%; Pred. No. 0.0027;
Matches 80; Conservative 55; Mismatches 98; Indels 79; Gaps 16;

QY 131 KNSLDVAVKHKSLIYENFDYILKRVPCVTSFLLDIYLFKIVYEIER--IEN- 186
DB 161 QSNKDKQSILRT-LFNSERFDEI-----RHLLVENVKQEK--VQIENRYTOIENL 208
QY 187 -HNFLFYKDESIVQLQDALSSATTSLSALTSQNNDRGSGILSSFLRKQSNHNSKDINSL 245
DB 209 WNDIDTFNNDELALYKELESSQTDKMIKPPQFNDYGCKILKSF--EAKNKITKELDDL 266
QY 246 RN---LNDLSQELARLKSNNLNEGMYTATPSASLEVIKYD-----LSYLKEALAL 294
DB 267 NHKYNVVELSENTKQLKA-----EKIKFDLLKQEQYIDKLKQELKM 309
QY 295 I---KAKIGADTKEPLTRS FNEQAKGLGNDGKDRSNYYDFLKGVOEVENACN----- 345
DB 310 IQESKVLITYFRLQSLKDKDELVSLEHQSKLNETHYNEIKGFQKQLEHLSTRENEIT 369

commercially important nucleic acid fragments of the *Staphylococcus* genome and/or plasmids; and identifying an expression modulating fragment of the *Staphylococcus* genome and/or plasmids. The methods and compositions of the present invention are useful for the diagnosis, prevention and/or treatment of an *Staphylococcal* epidermidis bacterial infection. This is the amino acid sequence of a *S. epidermidis* protein of the invention.

Sequence · 1010 AA;

Query Match	7.6%;	Score 157.5;	DB 8;	Length 1010;
Best Local Similarity	25.6%;	Pred. No. 0.0027;		
Matches 80;	Conservative	55;	Mismatches 98;	Indels 79;
				Gaps 16;

```

131 KNNSLDVAKTHKSRLIIYENPDYILKRYVPCYTESFLDDIVLPEKIVYVEIERR-----ISN- 186
162 QSNKSKQSDILRT-LFNSERFDEI-----RHLLVENVKQEK--VOIENRYTQIENL 209
187 -HNPLFYKDESLVQLQDALSSATTSLSALTQNNDRGSGIILSPFKQNSNNHKSQISNL 245
210 WNDIDTFNNDELALYKELESSQTDKMEKFPQFNDYGCKILKSF--EAKNKITKELDDL 267
246 RN---LNDLSQOELARUKSNLNEGFMFTAPSALEVIKYD-----LSYLKEALAL 294
268 NHKYKVNVELSENTRYKLUK-----EKIKFDDLUKKEQNYIDKLUKQELKM 310
295 I---KAKIGADTKPELTRSFNEOAKGLGNDCKGRSNIYDFLKGVBQEOVENACN----- 345
311 IQESKVLITFTRLQSLKQKQDELVSLEHQSKLNETNYHNEIKFQKQLEHLSTRENEIT 370
346 ----LKLTRYFGLDKMFNSL--IMLSBEQK--VERDIKLTLYSKYNQLI-----QS 389
371 QFNQYLEKNQVF-----FNQLDKLISSVQOKPVLIEEEK--RLTSEYNDLITKKEELTKE 423
390 SSFDNEELAILK 401
424 MNKNKDPFAIE 435

```

RESULT 7

ADA35569
ID ADA35569 standard; protein; 448 AA.

ADA35569:

AA	DT	20-NOV-2003	(first entry)

DE Acinetobacter baumannii protein #2730.

XX Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
KW plant biocontrol agent.
KW

XX OS Acinetobacter baumannii.

XX PN US6562958-B1.

XX
PD
13-MAY-2003.XX
PF 04-JUN-1999: 99US-00328352.XX
PR 09-JUN-1998: 98US-0088701P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX
PT
Breton G. : Bugh D.

XX
DB WPT: 2003-576092/54

DR WPI; 2003-576092/
DR N-PSDB: ADA31443.

New Acinetobacter baumannii proteins and nucleic acids, useful as reagents for diagnosing a bacterial disease, as components of antibacterial PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for plants.

XX
PS Example; SEQ ID NO 6856; 328pp; English.

CC The invention relates to isolated *Acinetobacter baumannii* nucleic acids.
CC The *A. baumannii* nucleic acids and polypeptides are useful as reagents
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of
CC *A. baumannii* and other *Acinetobacter* species in a sample, in screening
CC compounds for the ability to interfere with the *A. baumannii* life cycle
CC or to inhibit *A. baumannii* infection, and as biocontrol agents for
CC plants. The present sequence represents the amino acid sequence of an *A.*
CC *baumannii* protein.

Sequence 448 AA;

Query Match 7.4%; Score 153.5; DB 6; Length 448;
Best Local Similarity 21.8%; Pred. NO. 0.0018;
Matches 87; Conservative 59; Mismatches 145; Indels 109; Gaps 18;

[illegible]

RESULT 8

ADN19417

ID ADN19417 standard; protein; 1957 AA.

AC ADN19417;

DT 02-DEC-2004 (first entry)

Bacterial polypeptide #2070.

Recombinant DNA construct; transformed plant; improved plant property; cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis; pathogen tolerance; pest tolerance; plant disease resistance; cell cycle pathway modification; plant growth regulator; homologous recombination; seed oil yield; protein yield; carbohydrate; nitrogen; phosphorus; photosynthesis; lignin; galactomannan; bacterial polypeptide.

xx Bacteria.

XX
97086XX
FN
USZ003Z33675

FD-18-DEC-2003. XX

XX 21-FEB-2002; 2002US-0360039P.
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 2070; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 1957 AA;
SQ

Query Match 7.3%; Score 150.5; DB 8; Length 1957;
Best Local Similarity 20.5%; Pred. No. 0.024;
Matches 96; Conservative 82; Mismatches 183; Indels 107; Gaps 17;

QY 3 DLRKTKLIDKISLELYKYSIFP-----RVNIENVAEDCLKNGILLESAAHN 49
DB 697 DLRKSEALRFKLEAKNLREVIDNLKGHETLEAQRDLHLSLSDAKNTAILSELTK 756

QY 50 VSEVELARLKVO-----LKNALLNCIISYR-----FHGI--GYVLVTKT-DTLID 91
DB 757 SSE-DVKRLTANVETLTQSKAMKQSFSLVNSYQISINLYHELRDHDVWMSQNTLLE 815

QY 92 LEQPV-----NIEPIGFYLDYBYRDLGVDPFHITYKVNKNKNSLDAVKIH 141
DB 816 SESKLTDCENLTQQNTMLIDNVQKLMHKHVNQ-----ESKVSSELKVNKGLSLDLKNLR 870

QY 142 KSLRIIYENFDYIL-----KRYVPCYTESFLDYLFEKLYVEIERIENHNFPLYKD 194
DB 871 SLSLNAISNDQILTLQABLSEKNYDSLEQESQLNSGL-----KSLEAKQLLHTE 921

QY 195 ESIVLQDALSATTSLSALTQSNDRGSGILSSFLRKQNSNNHSDIENLNLNDSLSQ 254
DB 922 NE--ELHRLDKLTGKLEKSSDLGKILTA-----RQEISNLKEENWSQSQ 969

QY 255 ELARLKNLNNEGMFTTAFPSLASLEVIKYDLSYLVKALALIKAKIGADTKPELTRSFNBQ 314
DB 970 AITSVKSKLDE-----TLSSKSKLEA---DIEHLKNKVSEVEVE-----RNALLASNERL 1016

QY 315 AKGLGNDG-----KGRSNYYDFLKGVOEVENACNLKLTXYFGLDMKFNSL 361
DB 1017 MDDLKNGENIASLQTEIEKKAENDDLQSKLSVVSSEYENLLISSQTNKSLDKTNQL 1076

QY 362 IMLSE-----EQKVERDIKLIYLYSKYNQIQSSFDNEELAIKKEK 403
DB 1077 KYIEKNVQKLDEKQDNVEELTQYKGLGEENAIKQDELLALRKK 1124

RESULT 9
ADN18442
ID ADN18442 standard; protein; 1169 AA.
XX
AC ADN18442;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #1095.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 1095; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress

CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX SQ Sequence 1169 AA;
 Query Match 7.0%; Score 143.5; DB 8; Length 1169;
 Best Local Similarity 22.2%; Pred. No. 0.042;
 Matches 113; Conservative 87; Mismatches 167; Indels 141; Gaps 29;
 QY 4 LRKTKLIDKISSLELYKYSIFRNYIENVAEDCLKNGI-ILBSAHHNVSEV--LARLV 60
 Db 159 IERRKIIDEISGIAFDEK-----KKAEELKARELIMDIRISEVNNLKGLK 211
 QY 61 OLKNALLNCIIISYRFGHGVYLVKTKDILID--LEQPNVNIPLI-GFEVLDVEY---VRD 114
 Db 212 EKEDAEKYIKNEELKAAYAILKVKSYLVNLLR---NIQNDIKNLEKNEFLSKVRE 268
 QY 115 LGVDFDHTYKVKSNKNNKNSL-----DAVKIHKSLRIIYENFDYILKRYVPCYTESFLL 168
 Db 269 IDVEIENL--KLRLNINELNEKNEVEVLELKS-----IKLEVEIENDKKVL 316
 QY 169 D--IYLFKEIYVEIERR-----IENHNFYKDESIVQ-----LQ 201
 Db 317 DSSINELKKVEVEIENKKKEIKETQCKIENRDSIIIEKEQKETEKKINLVYKERLK 376
 QY 202 DALSSATYSLTSALTOSNDRGSGILSSFLRKNS-----NNHSDIISNRLND----- 250
 Db 377 EAIASESIIKHLKESEME-----IADETAKQNELYRLKLELNDLNLNKRNFIEKNN 432
 QY 251 ----SLSQELARLK-----SNLNMEGMEYATPSSASLEVIKYDL-----S 286
 Db 433 EMIKKLEETVEDVDYKPLYLELENLVE--IEFSKGIKELEKKELQAKDELHAE 491
 QY 287 YLKEALALIKA-----KIGADTKPELTRSFNEQAKGL-----GNDGK----- 324
 Db 492 YVKN-ARIKALKEMEELSMO--RAIREILNANLPGIIDIVGNLTKTKIEYKTAIEVAAG 548
 QY 325 DRSNYDFLKGVOEVENACNLKLTJK-----YFGLD-MKFNLSIMLSEEQKVERDIKLE 378
 Db 549 NRLNHI-VVKRMDDAVRAIKYLRKLGKGRATPLDREGREAYVIDEGVGIGRAIDLVE 607
 QY 379 LYSKNQLIQSSSPFN-----EELALKE 402
 Db 608 FDEYRRVFE-YVFGNTVVVENIDIAKE 634
 RESULT 10
 ABU19076
 ID ABU19076 standard; protein; 819 AA.
 XX AC ABU19076;
 XX DT 19-JUN-2003 (first entry)
 XX DE Protein encoded by Prokaryotic essential gene #4603.
 XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX OS Borrelia burgdorferi.
 XX PN WO200277183-A2.
 XX PD 03-OCT-2002.
 XX PF 21-MAR-2002; 2002MO-US009107.
 XX PR 21-MAR-2001; 2001US-00815242.
 XX PR 06-SEP-2001; 2001US-00948993.
 XX PR 25-OCT-2001; 2001US-0342923P.
 XX PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.
 XX PA (ELIT-) ELITRA PHARM INC.
 XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlson KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA22946.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 PS Claim 25; SEQ ID NO 47000; 1766pp; English.
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 819 AA;
 Query Match 6.8%; Score 141; DB 6; Length 819;
 Best Local Similarity 22.2%; Pred. No. 0.041;
 Matches 117; Conservative 93; Mismatches 152; Indels 164; Gaps 30;
 QY 4 LRK--TKLIDKISSLE--LYKYSIFRNYIENVAEDCLKNGIILBSAHHNV---EVELA 56
 Db 216 LRKHQDLSQKLEALEKITLTKLNFNFDLNV-----LRSELNLDGLSKFLSSGPKLE 271
 QY 57 RLKVLQKALLNCIIISYRFGHGVYLVKTKDILIDLEQPNVNIPLI-GFEVLDVEYVRDGL 116
 Db 272 FYSFREKVNN-----IQELSKDELMSKILGKLEIKQLEQKTKLKN 317
 QY 117 VDFDHTYKVKSNKNNKNSLDAVKIHKSLRIIYENFDYI-LKRYVPCYTESF--LDDIYLF 173
 Db 318 LEDKFKVKSSESEKNSINNLYQLNLSLKKNDPFLSDSEINRYTKSPFELVDLIL- 376
 QY 174 EKIVIEIRRIENHNFYKD---ESLVQLQDALS-----SATSLSALT 215
 Db 377 ----SVLRSAKSEFVLLKEHILDSKLFLSELSIKYIKIRVNLIKYIKDENLALL 431
 QY 216 QSNDRGSGI-----LSSFLRKONS--NNHSDIISNLR-----NLNDSLSQ----E 255
 Db 432 R---DKIEPIFDQNVNKKFLLEKLNALNSLAKELITNRIERLIEKTLQNLNDVIGLEXYE 488


```
XX SQ Sequence 642 AA;
Query Match 6.7%; Score 138; DB 2; Length 642;
Best Local Similarity 21.5%; Pred. No. 0.051;
Matches 104; Conservative 76; Mismatches 162; Indels 142; Gaps 25;

QY 24 PFRYNIENVAEDCLKNGILLESAAHNVSVELARLKVQLKNALLNCIIISYRPHGIGYVLV 83
Db 91 FIOQCIRQLCEFLTENG- - - - -AHNVSMKSLQAPSV--KDFLKIFTFLYGLFCLPSYELP 143
QY 84 KTKOTLIDLEQPNIELPFGFYLDY- - - - -VRDLGV--DPDHTYKVKSNKNNKSL 135
Db 144 DTK- - - - -FEE- - - - -EVPRIKDLGYPFALSKSSMYTVGAPHTWPHIVAALVW- - - - -LI 189
QY 136 DAVKIH- - - - -KSRLIYEN- - - - -FDYILKRYVPCYTESFLLDIYLF 174
Db 190 DCIKIHTAMKESPLFDDGQPWGEETEDGIMHKNLFLDYTIK- - - - -CY-ESFMSGADSF 244
QY 175 KIYVEIERRIENHNFLFYKDESILVOLQDALSSATTSLSALTQSNDRSGSILSSFLRK- - 232
Db 245 ENNAELQSKLD- - - - -LFNVDAFKLSLEAKNRALNEQIARLEQEREKEPNRLES-LRKLK 300
QY 233 - - - - -QNSNNHSDI- - - - -SNLRNLN- - - - -DSLQELARLKSNNNEG 267
Db 301 ASLQGDVQKYQAYMSNLSHSAILDQKLNGLNEETARVELECEETIKQENTRLQNIIDNQK 360
QY 268 MYTATPSASLEVIKYDLSYLKBALALIKAKIGADTKPTLRSFNEQAK-GLGNDG- - - - 322
Db 361 --YSV- - - - -ADIERINHNERNELOQTI- - - - -NKLTKOLEAQKLMNEELKYARGKEAIE 411
QY 323 - - - - -KG-DRSNYYDF- - - - - - - - - - -LKGQVQOVENA 343
Db 412 LAEYHKLARKLKLIPKGAENSKGYDPEIKFNPAGANCLVKYRAQVYVPLKELINETEE 471
QY 344 CNLKLTKYFGLDMKFNSL-IMLSEBOKVERDIK--LIELYSKYNQLIQSSSFDNBEALIL 400
Db 472 INKALNKKMGLEDTLLEQLNAMITESKRSVRTLKEEVQKLDLDLYQQKKEABEEDKCASE 531
QY 401 KEKL 404
Db 532 LESL 535

RESULT 14
ID ABU03917 standard; protein; 642 AA.
AC ABU03917;
XX 29-JAN-2003 (first entry)
XX Human expressed protein tag (EPT) #583.
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX Homo sapiens.
XX WO200278524-A2.
XX 10-OCT-2002.
XX 28-MAR-2002; 2002WO-US009671.
XX 28-MAR-2001; 2001US-0279495P.
XX 21-MAY-2001; 2001US-0292544P.
XX 08-AUG-2001; 2001US-0310801P.
XX 01-OCT-2001; 2001US-0326370P.
XX 04-DEC-2001; 2001US-0336780P.
```

```
PR 20-FEB-2002; 2002US-0358985P.
XX (ZYCO-) ZYCOS INC.
XX Chiciz RM, Tomlinson AJ, Urban RG;
XX WPI; 2003-040607/03.
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
PS Example 2; SEQ ID NO 583; 134pp; English.
XX The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 642 AA;
```

```
Query Match 6.7%; Score 138; DB 6; Length 642;
Best Local Similarity 21.5%; Pred. No. 0.051;
Matches 104; Conservative 76; Mismatches 162; Indels 142; Gaps 25;

QY 24 PFRYNIENVAEDCLKNGILLESAAHNVSVELARLKVQLKNALLNCIIISYRPHGIGYVLV 83
Db 91 FIOQCIRQLCEFLTENG- - - - -AHNVSMKSLQAPSV--KDFLKIFTFLYGLFCLPSYELP 143
QY 84 KTKOTLIDLEQPNIELPFGFYLDY- - - - -VRDLGV--DPDHTYKVKSNKNNKSL 135
Db 144 DTK- - - - -FEE- - - - -EVPRIKDLGYPFALSKSSMYTVGAPHTWPHIVAALVW- - - - -LI 189
QY 136 DAVKIH- - - - -KSRLIYEN- - - - -FDYILKRYVPCYTESFLLDIYLF 174
Db 190 DCIKIHTAMKESPLFDDGQPWGEETEDGIMHKNLFLDYTIK- - - - -CY-ESFMSGADSF 244
QY 175 KIYVEIERRIENHNFLFYKDESILVOLQDALSSATTSLSALTQSNDRSGSILSSFLRK- - 232
Db 245 ENNAELQSKLD- - - - -LFNVDAFKLSLEAKNRALNEQIARLEQEREKEPNRLES-LRKLK 300
QY 233 - - - - -QNSNNHSDI- - - - -SNLRNLN- - - - -DSLQELARLKSNNNEG 267
Db 301 ASLQGDVQKYQAYMSNLSHSAILDQKLNGLNEETARVELECEETIKQENTRLQNIIDNQK 360
QY 268 MYTATPSASLEVIKYDLSYLKBALALIKAKIGADTKPTLRSFNEQAK-GLGNDG- - - - 322
Db 361 --YSV- - - - -ADIERINHNERNELOQTI- - - - -NKLTKOLEAQKLMNEELKYARGKEAIE 411
QY 323 - - - - -KG-DRSNYYDF- - - - - - - - - - -LKGQVQOVENA 343
Db 412 LAEYHKLARKLKLIPKGAENSKGYDPEIKFNPAGANCLVKYRAQVYVPLKELINETEE 471
QY 344 CNLKLTKYFGLDMKFNSL-IMLSEBOKVERDIK--LIELYSKYNQLIQSSSFDNBEALIL 400
Db 472 INKALNKKMGLEDTLLEQLNAMITESKRSVRTLKEEVQKLDLDLYQQKKEABEEDKCASE 531
QY 401 KEKL 404
```


Db 532 LESL 535

RESULT 15

ABU03918

ID ABU03918 standard; protein; 642 AA.

XX

AC ABU03918;

XX

DT 29-JAN-2003 (first entry)

XX

XX Human expressed protein.tag (EPT) #584.

XX

DE Translational profiling; expressed protein tag; EPT; kinase; phosphatase;

XX protease; protease inhibitor; transporter; cytoskeletal protein;

KW receptor; transcription factor; cancer; MHC;

KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;

XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX

OS Homo sapiens.

XX

PN WO200278524-A2.

XX

XX 10-OCT-2002.

XX

XX 28-MAR-2002; 2002WO-US009671.

PF

XX 28-MAR-2001; 2001US-0279495P.

PR

XX 21-MAY-2001; 2001US-0292544P.

PR

XX 08-AUG-2001; 2001US-0310801P.

PR

XX 01-OCT-2001; 2001US-0328370P.

PR

XX 04-DEC-2001; 2001US-0336780P.

PR

XX 20-FEB-2002; 2002US-0358985P.

XX

XX (ZYCO-) ZYCOS INC.

XX

XX Chicx RM, Tomlinson AJ, Urban RG;

PI

XX WPI; 2003-040607/03.

DR

XX

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,

PT cytoskeletal proteins, receptors or transcription factors), useful for

PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or

PT leukemia.

XX

XX Example 2; SEQ ID NO 584; 134pp; English.

PS

XX

XX The invention describes a purified polypeptide, which comprises a

CC fragment of a kinase, phosphatase, protease, protease inhibitor,

CC transporter, cytoskeletal protein, receptor or transcription factor. The

CC polypeptide is useful as an immunogenic composition for eliciting in a

CC mammal an immunogenic response directed against any of the purified

CC polypeptide. The purified polypeptide, or the antibody that binds to this

CC polypeptide, is useful for treating cancer. The polypeptide is also

CC useful for identifying compounds that binds to a naturally processed

CC class I or class II MHC-binding polypeptide. The polypeptides and

CC polynucleotides are particularly useful for treating or preventing

CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,

CC lymphoma or leukaemia. These are also useful for screening agents for

CC treating the above mentioned diseases. This sequence represents an

CC expressed protein tag (EPT) isolated from human tissue for translational

CC profiling. Note: This sequence does not appear in the printed

CC specification but was obtained in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX

XX Sequence 642 AA;

SQ

Query Match 6.7%; Score 138; DB 6; Length 642;

Best Local Similarity 21.5%; Pred. No. 0.051;

Matches 104; Conservative 76; Mismatches 162; Indels 142; Gaps 25;

QY 24 FFRNYIENVAEDCLKNGLILESAAHNVSEVELARLVQLKALLNCIISYRPHGIGYLV 83

Search completed: January 24, 2006, 19:44:26

Job time : 93.9538 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006. Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:57:15 ; Search time 71.5326 Seconds
(without alignments)
2377.329 Million cell updates/sec

Title: US-10-688-058-4
Perfect score: 2062
Sequence: 1 MCDLRKTKLIDKISSLELYK.....QSSSFDNEELAIKKEKLFSP 407

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA_Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pcp.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2062	100.0	407	5	US-10-688-058-4
2	566	27.4	119	5	US-10-688-058-2
3	157.5	7.6	1009	4	US-10-282-122A-71141
4	157.5	7.6	1010	4	US-10-724-972A-4177
5	150.5	7.3	1957	4	US-10-369-493-2070
6	150.5	7.3	1957	5	US-10-732-923-3328
7	143.5	7.0	1169	4	US-10-369-493-1095
8	141.5	6.9	2910	5	US-10-732-923-3342
9	141	6.8	819	4	US-10-282-122A-47000
10	141	6.8	1527	5	US-10-732-923-3354
11	140.5	6.8	864	5	US-10-732-923-3312
12	140.5	6.8	921	5	US-10-732-923-3305
13	140	6.8	703	4	US-10-282-122A-63529
14	138	6.7	642	4	US-10-648-593-172
15	138	6.7	642	4	US-10-755-889-608
16	138	6.7	642	5	US-10-733-878-441
17	138	6.7	642	5	US-10-473-127-583
18	138	6.7	642	5	US-10-473-127-584
19	138	6.7	642	5	US-10-473-127-585
20	138	6.7	642	5	US-10-473-127-587
21	138	6.7	642	5	US-10-473-127-588
22	138	6.7	642	5	US-10-287-436A-557
23	138	6.7	664	5	US-10-500-530-50
24	135.5	6.6	696	4	US-10-408-765A-1067
25	134	6.5	1939	5	US-10-732-923-3340
26	133.5	6.5	2503	5	US-10-828-985A-11
27	133.5	6.5	2543	5	US-10-828-985A-9

28 133.5 6.5 2568 5 US-10-828-985A-7 Sequence 7, Appli
29 133.5 6.5 2633 5 US-10-450-763-36864 Sequence 36864, A
30 133.5 6.5 2663 5 US-10-723-860-749 Sequence 749, App
31 131.5 6.4 878 4 US-10-724-972A-4767 Sequence 4767, Ap
32 131.5 6.4 1184 4 US-10-282-122A-53254 Sequence 53254, A
33 129.5 6.3 365 4 US-10-125-692-16 Sequence 16, Appl
34 129.5 6.3 365 5 US-10-991-347-16 Sequence 16, Appl
35 129.5 6.3 575 4 US-10-601-036-8 Sequence 8, Appli
36 129.5 6.3 1182 4 US-10-282-122A-53445 Sequence 53445, A
37 128.5 6.2 996 3 US-09-815-242-5251 Sequence 5251, Ap
38 128.5 6.2 1009 3 US-09-815-242-12141 Sequence 12141, A
39 128.5 6.2 1009 4 US-10-282-122A-43832 Sequence 43832, A
40 126 6.1 980 5 US-10-732-923-3343 Sequence 3343, Ap
41 126 6.1 2133 5 US-10-732-923-15030 Sequence 15030, A
42 125 6.1 436 4 US-10-120-801-88 Sequence 68, Appl
43 125 6.1 436 4 US-10-275-595A-16 Sequence 16, Appl
44 125 6.1 1478 5 US-10-732-923-3353 Sequence 3353, Ap
45 125 6.1 1875 4 US-10-369-493-22285 Sequence 22285, A

ALIGNMENTS

RESULT 1

US-10-688-058-4

; Sequence 4, Application US/10688058

; Publication No. US20050058661A1

; GENERAL INFORMATION:

; APPLICANT: SYKES, KATHRYN F.

; APPLICANT: HALE, KATHERINE S.

; APPLICANT: JOHNSTON, STEPHEN A.

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING

; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF THE GENUS

; TITLE OF INVENTION: BORRELIA

; FILE REFERENCE: MCRO:00305

; CURRENT APPLICATION NUMBER: US/10/688,058

; CURRENT FILING DATE: 2003-10-17

; NUMBER OF SEQ ID NOS: 141

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 4

; LENGTH: 407

; TYPE: PRT

; ORGANISM: Borrelia burgdorferi

; US-10-688-058-4

Query Match 100.0%; Score 2062; DB 5; Length 407;
Best Local Similarity 100.0%; Pred. No. 5.8e-149;
Matches 407; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCDLRKTKLIDKISSLELYKYSIFPRNYIENVAEDCLKNGLILESAAHNVSEVELARLKV 60
Db 1 MCDLRKTKLIDKISSLELYKYSIFPRNYIENVAEDCLKNGLILESAAHNVSEVELARLKV 60

QY 61 QLKXALLNCIIISYRPHGIGYLVKTKOTLIDLEQPNIELPFGFYLYDYVYVDLGVDPD 120
Db 61 QLKXALLNCIIISYRPHGIGYLVKTKOTLIDLEQPNIELPFGFYLYDYVYVDLGVDPD 120

QY 121 HITVYKNNKNNSLDAVKIHKSRLLIYENFDYILKRVPCVTSFLLDIYLFKIIYVEI 180
Db 121 HITVYKNNKNNSLDAVKIHKSRLLIYENFDYILKRVPCVTSFLLDIYLFKIIYVEI 180

QY 181 ERIENHNFLYKDESLVQLQDALSSATTSLSALTQSNDRGSGILSSFLRKQNNHNSK 240
Db 181 ERIENHNFLYKDESLVQLQDALSSATTSLSALTQSNDRGSGILSSFLRKQNNHNSK 240

QY 241 DISNLRNLNLSLSQELARLKNLNNEGFFYATPSASLEVIKYDLSYLKEALALIKAG 300
Db 241 DISNLRNLNLSLSQELARLKNLNNEGFFYATPSASLEVIKYDLSYLKEALALIKAG 300

QY 301 ADTKEPLTRSFNEQAKGLGNDGKDRSNYDFLKGQVQOVENACNLKLTKYFGLDMKFN 360
Db 301 ADTKEPLTRSFNEQAKGLGNDGKDRSNYDFLKGQVQOVENACNLKLTKYFGLDMKFN 360


```

RESULT 7
US-10-369-493-1095
; Sequence 1095, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1095
; LENGTH: 1169
; TYPE: PRT
; ORGANISM: Methanococcus jannaschii
US-10-369-493-1095

Query Match      7.0%; Score 143.5; DB 4; Length 1169;
Best Local Similarity 22.2%; Pred. No. 0.063;
Matches 113; Conservative 87; Mismatches 167; Indels 141; Gaps 29;

QY 4 LRTKLIIDKISLELYKYISIFRNVIENVAEDCLKNGL-IIESAAHNVSEVE--IARLKV 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 159 TERRKIIDEISGIAFDEK-----KKAAEELKARELIMIDIRISEVENNLKLLK 211
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 OLKNALLNCIIISYRFGIGYLVKTDLTID--LQPVNIELPI-GPEVLDVEY---VRD 114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 212 EKEDAEKIKUNEELKAAYALILKVSYNLLE---NIQNDIKNLEELKNEFLSKVRE 268
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 115 LGVDFDHTYKYVKNKNNKSL-----DAVKIHSRLIYENFDYILKRYVPCYTESPLL 168
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 269 IDVEIENL--KLRLNNIINELNEKNEEVLHLKS-----IKELEVEIENDKKVL 316
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 169 D--IYLFKRYVVEIERR-----IENHFLFYKDESIVQ-----LQ 201
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 317 DSSINELKKEVEIEENKKKEIKETOKKIENRDSIIEKEQIKETEKKIKNLNVEKERLK 376
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 202 DALSSATTSLSALTOSNDRSGIILSSFLRKNS-----NNHSKOISNLRLND----- 250
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 377 EAIASESIIHLKESME-----IADETAKNQNELYRLKLELNDLDNLNRKNFIEKNN 432
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 251 ----SLSQELARLK-----SNLNNMGFYTATPSASLEVIKYDL-----S 286
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 433 EMIKKLKSELEIIVEDVDTKPLYLELENLVE-IFFSKRGIKLEKKEKQKLDLHAE 491
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 287 YLKEALAIKA-----KITGADTKPELTFRSFNOAKGL-----GNQDK-----G 324
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 492 YVKN-ARIKALKEMEELSM-D--RAIRELTNANLPGIIDIIVGNLKGTKIEVKTATEVAAG 548
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 325 DRSNYVDFLKGVOEVENACMLKTK-----YFGLD-MKFNSLIMLSEOKVERDIKLE 378
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 549 NRLNHI-VYKRMDDAVRAIKYKLERKLGKRGATPLDRIEGREAYYIDEDGVIGRAIDLVE 607
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 379 LYSKYNQLIQSSSFON-----BELAILKE 402
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 608 PDEKTRVPE-YVFGNTVVVENIDIKE 634
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
US-10-732-923-3342
; Sequence 3342, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C

```

```

; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3342
; LENGTH: 2910
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-732-923-3342

Query Match      6.9%; Score 141.5; DB 5; Length 2910;
Best Local Similarity 19.2%; Pred. No. 0.3;
Matches 81; Conservative 83; Mismatches 162; Indels 95; Gaps 18;

QY 11 DKISLELYKYISIFRNVIENVAEDCLKNGLILSAAHNVSEVELARUKVLKNALLNCI 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 299 NNIDNKELY---YFNESINEIHENCTDSKILKILKICNIIFE-----KKTKN----- 342
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 71 ISYRFGIGYLVKTDLTIDLEQPVNIELPIGFEYLDVEYVRDLGVDFDHTYKYVKN 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 343 ----IH-----ILKQLDLLFLQK-----SYENRNIFYEKKSLYGMENISSKDNINN 385
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 131 KNNSLDAVKIHSRLIYENFDYILKRYVPCYTESFLDDIYLFKRYVVEIERRIENHPL 190
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 386 DNNNNN-----YHQNFSFGKENYVFLDNNNTSINIVREDE---SKEKSI SNYF- 432
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 191 FYKDESLVOLQDALSSATTSLSALTQSND-----RSGSILSSFLRKQNSNHSKDISNLR 246
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 433 ---DKTLFLFNVLKVSHTISGLYNTFMDLTNYSNDSALNNHINSNNNNNNNNNNNN 489
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 247 NLND-SLSQELARLKNLNNMGFYTATPSASLEVIKYDLSYK-----EALALIKAKI 299
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 490 NNDYFISYEURLKIESIN-----YQRAVNAALYLFHILHMLLETTSOKNNMLIENEI 545
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 300 GADTKPELTFRSFE-----QAKGLGNDGKGRSNNYDFLKGVOEQVE----- 341
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 546 --STIRNFIKINKKKVDSNININIVQYKNVPTQNNNTYKSNKQLNIQNNENKKYV 603
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 342 NACNLKTYFGLDMKNSLIMLSEOKVERDIK-----LIELYKYNQLIQSSSFONE 395
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 604 NLINKELVRH---NFEIRKIIICKYNNKIIKWMKKIHNLYIMHYVYSYKN--DQNGYINE 658
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 396 E 396
DB 659 E 659

RESULT 9
US-10-282-122A-47000
; Sequence 47000, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23

```


Query Match 6.8%; Score 140; DB 4; Length 703;
Best Local Similarity 21.3%; Pred. No. 0.059;
Matches 105; Conservative 73; Mismatches 140; Indels 176; Gaps 26;

QY 4 LRKTKLIDKISSLELYKYKSIFFRNYIENVAEDCLKNGILLESAAHNVSEVELARLVQV-- 61
DB 94 LKRFILIDD-----DISYLFKEPLADSKLDINLAKQIIDFNKTFPADFEINKLDQDER 146
QY 62 -----LKNALLNCIIISYR-----FHGIGYVVLVKT----- 85
DB 147 LISLCEHSLNKNDBEYSTLKTQLINAFISYEKNKILNKNLDFHDL---LIKTCNLLSNDN 203
QY 86 -----KDTLIDLEQPNIELPIGFEYLDVEYVVDLGVDPDHITYKVKSNK----- 131
DB 204 DLLNQMSSQFOHILVDFQDTN-----QIQVELIKWL-VTRKNLFLVGDNNQMIYR 254
QY 132 -----NNSLDAVK-----THKSLRIIYENFDYILKRYVPCYTESFLLDIYLPK----- 175
DB 255 WRGAVNGIITALKHDFNVPKSN-BFFINQNYRCQDQNILAVANQIILLIMAYEKQVKTEKN 313
QY 176 -----IYVEIERRIENH-NFLFYKDESILVQDALSSATTSLTSALTQSNNDRG 222
DB 314 LLFSTLNGDKXPVYQAE-SVENQANWIFNKIKALNQ-----TEKINPKD 357
QY 223 SGILSSFLRKQSNHNSKDIENLRLNDSLSQ-ELARLKSNNLNEGMFYTATPSASLEVI 281
DB 358 MAIL-----FRK-----NRDITTMVELIEADGTIPLPKQKSYFN-----QLV 394
QY 282 KYDLSYLKEALALIKAKIGADTKETPLTRSFNEQAKGLNDGK-----GDRSNYYDFLKG 336
DB 395 K-----LQRLVIAISTRTNLDIKRAL-----QALKIWSNDLKELWKQSKDKNLFPDLKWS 444
QY 337 Q-EQVENACNLKTKYFGLDMKFNLSIMLSEBQKVERDIKLIYLS--KYNQLIQSSSFD 393
DB 445 ELNQKHSSKLKATGY-----FNLLIKLAEDQOI--NLLPTELKFLKLVQDTIEN---- 492
QY 394 NEELAILKEKLFSP 407
DB 493 -----LLWKKLGTET 501

RESULT 14
US-10-648-593-172
; Sequence 172, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 172
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-648-593-172

Query Match 6.7%; Score 138; DB 4; Length 642;
Best Local Similarity 21.5%; Pred. No. 0.075;
Matches 104; Conservative 76; Mismatches 162; Indels 142; Gaps 25;

QY 24 FFRNYIENVAEDCLKNGILLESAAHNVSEVELARLVQKNALLNCIIISYRPHGIGYVLV 83
DB 91 FIOQCIRQLCEPLTENG-----AHNVSMKSLQAPSV--KDFLKIFTFLYGLCPSYELP 143
QY 84 KTKDTLIDLEQPNIELPIGFEYLDYDEY-----VRDLGV--DFDHITYKVKSNKNNKNSL 135
DB 144 DTK-----FEE-----EVPRIKDLGYPPALSKSSMYTVGAPHTWPHIVAALVW-----LI 189
QY 136 DAVKIH---KSLRIIYEN-----FYILKRYVPCYTESFLLDIYLF 174
DB 190 DCIKIHTAMKSSPLFDDGQPGWGETEDGIMHNKLFIDYTIK---CY-ESFMSGADSF 244
QY 175 KIYVEIERRIENHNFYKDESILVQDALSSATTSLTSALTQSNNDRGSGILSSFLRK-- 232
DB 245 EMNAELQSKLDK---LFNVDAFKLESLEAKNRALNEQIARLEQEREKEPNRLES--LRKLK 300

Query Match 6.7%; Score 138; DB 4; Length 642;
Best Local Similarity 21.5%; Pred. No. 0.075;
Matches 104; Conservative 76; Mismatches 162; Indels 142; Gaps 25;

QY 24 FFRNYIENVAEDCLKNGILLESAAHNVSEVELARLVQKNALLNCIIISYRPHGIGYVLV 83
DB 91 FIOQCIRQLCEPLTENG-----AHNVSMKSLQAPSV--KDFLKIFTFLYGLCPSYELP 143
QY 84 KTKDTLIDLEQPNIELPIGFEYLDYDEY-----VRDLGV--DFDHITYKVKSNKNNKNSL 135
DB 144 DTK-----FEE-----EVPRIKDLGYPPALSKSSMYTVGAPHTWPHIVAALVW-----LI 189
QY 136 DAVKIH---KSLRIIYEN-----FYILKRYVPCYTESFLLDIYLF 174
DB 190 DCIKIHTAMKSSPLFDDGQPGWGETEDGIMHNKLFIDYTIK---CY-ESFMSGADSF 244
QY 175 KIYVEIERRIENHNFYKDESILVQDALSSATTSLTSALTQSNNDRGSGILSSFLRK-- 232
DB 245 EMNAELQSKLDK---LFNVDAFKLESLEAKNRALNEQIARLEQEREKEPNRLES--LRKLK 300

DB 144 DTK-----FEE-----EVPRIKDLGYPPALSKSSMYTVGAPHTWPHIVAALVW-----LI 189
QY 136 DAVKIH---KSLRIIYEN-----FYILKRYVPCYTESFLLDIYLF 174
DB 190 DCIKIHTAMKSSPLFDDGQPGWGETEDGIMHNKLFIDYTIK---CY-ESFMSGADSF 244
QY 175 KIYVEIERRIENHNFYKDESILVQDALSSATTSLTSALTQSNNDRGSGILSSFLRK-- 232
DB 245 EMNAELQSKLDK---LFNVDAFKLESLEAKNRALNEQIARLEQEREKEPNRLES--LRKLK 300
QY 233 -----QSNNNHNSKOI-SMLRLNL-----DLSLSQELARLKSNNLNEG 267
DB 301 ASLOQDVQKYQAYMSNLSESAILLDQKLNGLNBEIARVELECEETIKQENTRLQNIIDNQK 360
QY 268 MFTYATTSASLEVIKYDLSYLKEALALIKAKIGADTKETPLTRSFNEQAK-GLGNDG---- 322
DB 361 --YSV--ADIERINHNERNELQQTII---NKLTKDLAEQOKLWNEELKYARGKEAIBTQ 411
QY 323 -----KG-DRSNYYDF-----LKGVOEQVENA 343
DB 412 LAEYHKLARKLKLIPKGAENSGYDFEIKFNPAGANCLVKYRAQVYVPLKELLNETSEE 471
QY 344 CNLKLTKYFGLDMKFNLSIMLSEBQKVERDIK-LIELYSKYNQLIQSSSFDNBEILAIL 400
DB 472 INKALNKKMGLEDLEQLNAMIETESKRSVRLKEEVQKLDLDLYQOKIKEAEEDEKCASE 531
QY 401 KEKL 404
DB 532 LESL 535

RESULT 15
US-10-755-889-608
; Sequence 608, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 608
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-755-889-608

Query Match 6.7%; Score 138; DB 4; Length 642;
Best Local Similarity 21.5%; Pred. No. 0.075;
Matches 104; Conservative 76; Mismatches 162; Indels 142; Gaps 25;

QY 24 FFRNYIENVAEDCLKNGILLESAAHNVSEVELARLVQKNALLNCIIISYRPHGIGYVLV 83
DB 91 FIOQCIRQLCEPLTENG-----AHNVSMKSLQAPSV--KDFLKIFTFLYGLCPSYELP 143
QY 84 KTKDTLIDLEQPNIELPIGFEYLDYDEY-----VRDLGV--DFDHITYKVKSNKNNKNSL 135
DB 144 DTK-----FEE-----EVPRIKDLGYPPALSKSSMYTVGAPHTWPHIVAALVW-----LI 189
QY 136 DAVKIH---KSLRIIYEN-----FYILKRYVPCYTESFLLDIYLF 174
DB 190 DCIKIHTAMKSSPLFDDGQPGWGETEDGIMHNKLFIDYTIK---CY-ESFMSGADSF 244
QY 175 KIYVEIERRIENHNFYKDESILVQDALSSATTSLTSALTQSNNDRGSGILSSFLRK-- 232
DB 245 EMNAELQSKLDK---LFNVDAFKLESLEAKNRALNEQIARLEQEREKEPNRLES--LRKLK 300

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:59:25 ; Search time 6.51393 Seconds
(without alignments)
633.203 Million cell updates/sec

Title: US-10-688-058-4
Perfect score: 2062
Sequence: 1 MCDLRKTKLIDKISSELYK.....QSSSFDNELAILKELFSF 407

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 70609 seqs, 10134256 residues

Total number of hits satisfying chosen parameters: 70609

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:
1: /cgn2_6/protdata/1/pubpaa/US08_NEW_PUB.pap:
2: /cgn2_6/protdata/1/pubpaa/US06_NEW_PUB.pap:
3: /cgn2_6/protdata/1/pubpaa/US07_NEW_PUB.pap:
4: /cgn2_6/protdata/1/pubpaa/US09_NEW_PUB.pap:
5: /cgn2_6/protdata/1/pubpaa/US10_NEW_PUB.pap:
6: /cgn2_6/protdata/1/pubpaa/US11_NEW_PUB.pap:
7: /cgn2_6/protdata/1/pubpaa/US12_NEW_PUB.pap:
8: /cgn2_6/protdata/1/pubpaa/US60_NEW_PUB.pap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	157.5	7.6	593	6	US-10-793-626-1358
2	135.5	6.6	2665	7	US-11-124-368A-214
3	135.5	6.6	2668	7	US-11-124-368A-215
4	126	6.1	5024	6	US-10-793-626-2964
5	114	5.5	710	6	US-10-467-657-4292
6	113.5	5.5	776	7	US-11-033-039-637
7	113.5	5.5	989	6	US-10-821-234-975
8	112	5.4	1167	7	US-11-052-554A-121
9	111	5.4	976	7	US-11-155-288-20
10	107.5	5.2	888	7	US-11-077-550-112
11	106.5	5.2	1078	7	US-11-165-221-43
12	106.5	5.2	1078	7	US-11-165-226-53
13	106	5.1	2710	7	US-11-051-453-41
14	105.5	5.1	1732	6	US-10-055-877-147
15	104	5.0	698	6	US-10-793-626-2388
16	102.5	5.0	864	7	US-11-077-550-102
17	102.5	5.0	865	7	US-11-077-550-100
18	102.5	5.0	866	7	US-11-077-550-104
19	102.5	5.0	867	7	US-11-077-550-80
20	102.5	5.0	876	7	US-11-077-550-82
21	102.5	5.0	1420	7	US-11-077-550-110
22	102.5	5.0	1614	6	US-10-821-234-903
23	101.5	4.9	605	6	US-10-689-742-140
24	101	4.9	1189	7	US-11-074-176-134
25	101	4.9	1735	7	US-11-040-472-13

26	100.5	4.9	1145	6	US-10-793-626-1432	Sequence 1432, Ap
27	100	4.8	465	6	US-10-878-556A-116	Sequence 116, App
28	100	4.8	495	6	US-10-821-234-1154	Sequence 1154, Ap
29	99	4.8	1095	6	US-10-793-626-3154	Sequence 3154, Ap
30	98.5	4.8	1216	6	US-10-873-528-12	Sequence 12, Appl
31	97.5	4.7	404	6	US-10-793-626-1130	Sequence 1130, Ap
32	97.5	4.7	860	7	US-11-077-550-175	Sequence 175, App
33	97.5	4.7	862	7	US-11-077-550-94	Sequence 94, Appl
34	97.5	4.7	866	7	US-11-077-550-88	Sequence 88, Appl
35	97.5	4.7	867	7	US-11-077-550-96	Sequence 96, Appl
36	97.5	4.7	867	7	US-11-077-550-98	Sequence 98, Appl
37	97.5	4.7	870	7	US-11-077-550-92	Sequence 92, Appl
38	97.5	4.7	871	7	US-11-077-550-84	Sequence 84, Appl
39	97.5	4.7	871	7	US-11-077-550-86	Sequence 86, Appl
40	97.5	4.7	871	7	US-11-077-550-90	Sequence 90, Appl
41	97.5	4.7	1070	7	US/11/062	Sequence 4, Appli
42	97.5	4.7	1095	7	US/11/062	Sequence 7, Appli
43	97.5	4.7	1169	7	US-11-077-550-20	Sequence 20, Appl
44	97	4.7	299	7	US-11-156-084-288	Sequence 288, App
45	97	4.7	568	6	US-10-793-626-2482	Sequence 2482, Ap

ALIGNMENTS

RESULT 1
US-10-793-626-1358
; Sequence 1358, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIORITY FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1358
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-1358

Query Match	7.6%	Score 157.5;	DB 6;	Length 583;
Best Local Similarity	25.6%	Pred. No. 0.00046;		
Matches	80;	Conservative 55;	Mismatches 98;	Indels 79; Gaps 16
QY	131	KNNSLDVAKHKSRLLIYENFDYILKRYVPCYTESFLDLIDYLFEXIYVEIERR-----IEN-186		
Db	154	QNSKDKQSILRT-LFNSERFDEI-----RHLLVENVKQEK--VQIENRYQIENL 201		
QY	187	-HNFIFYKDESIVQIQDLSATTSLSALTQSNDRGSGILSSFLRKQNNHSDINL 245		
Db	202	WNIDITFNDELALYKELESSQTKWIEKFPQFNDYGCKILKSF--BEAKNKITTELDDL 259		
QY	246	RN-----LNDSLSQELARLKSNLNNEGMFTATPSASLEVIKYD-----LSYLKEALAL 294		
Db	260	NHKYKVNVELSENTKLLKA-----EKIKFDLKKQBNYIDKLQELCKM 302		
QY	295	I---KAKIGADTKPELTRSFNEQAKGLGNDGDRSNYYDFLKGQVQEVENACN-----345		
Db	303	IQESKVLITYFTRLOSLKDKDELVSILHQSKLNETHYNEIKGQKQLEHLSTRENEIT 362		
QY	346	-----LKLTKYFGLDMKFNLSL--IMLSBEQ--VERDIKLIELYSKYNQI-----QS 389		
Db	363	QFNQVLEKNQVF-----FNQDKLISSVQKQKPVIEEEK--RLYSEYNDLITKCELTKE 415		
QY	390	SSFDNEELAILK 401		

Db 61 KLEKVPSPVLEMYK-AIGGKIYI--VDGDITKH-ISLEALSEDKKKIK-----DIYKDA 112
QY 66 LANCIIISYRFHGIGVYLV--KTKDTLIDLLEQPVNTELPIG-----FEYLDYEVYRDLG 116
Db 113 LUHEHYVAKGEBPVLIQSSSEYVTEKALNVYVIGKILSDILSKINQPKFLD 172
QY 117 V-----DFDH-----ITYKVKNNKNSLDVAKIHKSL--IYENFDY-----I 154
Db 173 VLNTIKNASDSQDQLLFTNQLKEHPTDFSVEFFLEQNSNEVEQVAKAPAYIEPQHRDV 232
QY 155 LKRYVPCYTESP-LLDIYLFYKIVYIEIRIENHNFPLFYKDBSLVOLQDASSATTSLSA 213
Db 233 LQLYAP--EAFNYMDKFEQINLSLEB-----LKDQMLSRYEKWEKIKQHYQH 280
QY 214 LTQSNDRSGSL-----SSFLRKQNSNNHSHK 240
Db 281 WSDSLSESGRLKKLQIPIEPKDDIHSLSQBEKELLKRIQIDSDFL-----STERKE 336
QY 241 DISNLR-NLNDLSQBE-----LARKNSLN--NEGMPFYTA 272
Db 337 FLKQLQIDRDSLSBEKELLNRIOVDSNPLSEKEKEFLKLDIOPIDINQRLQDTG 396
QY 273 -----TPSASLEVIK-----YDLSYLKEALAL--IKAKIGADTYKE 305
Db 397 GLIDSPSINLVRKQYKRDQIWNALLHQSIGSTLYNKIYLYENNINNLNLTATLGADLVD 456
QY 306 PLTRS-----FNEQAKGLGNDGKGRNYYDFLKGVOEQ----- 339
Db 457 STDNTKINRGIFNEFKK--NFKYSISNY--MIVDINERPALDNERLKWRQLSPDTRA 511
QY 340 --VENACNLKLTYPGLDMKFNLSMLSEEQKVERDIKLIELYSKYNQLOIQSSSPD---- 393
Db 512 GYLENG-KLILQNRNIGLEIKDVOILKQSEKEVIRIDAKVVP-KSKIDTKIQEAQLNINQE 569
QY 394 -NEELAILK-EKLPSF 407
Db 570 WNKALGLPKYTKLITF 585

RESULT 7
US-10-821-234-975
; Sequence 975, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 975--
; LENGTH: 989
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-975

Query Match 5.5%; Score 113.5; DB 6; Length 989;
Best Local Similarity 20.8%; Pred. No. 1.1;
Matches 98; Conservative 80; Mismatches 158; Indels 135; Gaps 22;
QY 29 IENAECDCKNGLILESAAHNVSEVELARLKVQLKN-----ALLNCIIISYRFHGIGVYL 82
Db 443 LQEIQLDQKR--LES-----SEAEKQLQVELQSRRAELVCLNTEISENSSDLSQKL 494
QY 83 VKTKD-----TLIDLQPVNTELPIGFEYLD-YEYVRDLGVDFDHITYKYKSNKNNS 134
Db 495 KETOSKYBEAMKEVLSVQKMKLGL-VSPESMDNYSHFELAVTBEET--NVLKQDLQNA 551

QY 135 LDAVKIHKSLR-----LIYENFDYILKRYVPCYTESFLLDIYLFYKIVYIEIRIENHN 188
Db 552 LEESERNKREVBLEBKLVREKGTVIKPPVEYEE-----MKSSYCSV---IENNN 600
QY 189 ----FLFYK-----DESLVOLQDASSATTSLSALTQSNDRGSGITLSSFLRKQNSNNHSHK 240
Db 601 KEKAFLEKYOEAQEEIMKLDTLK-----SQMTQASDEADMEKAMRMIDELN--K 652
QY 241 DISNLRNL-----NDSLSQ-- 254
Db 653 QVSELSQLYKEAQAELEDYRKRSKLESDVTAHYIHKABHEKMLQMTNVSVRAKAEDALSEMK 712
QY 255 -----ELARKNSLN-----NEGMPFYTATPSASLEVIKYDLSYLKEALALIKA 297
Db 713 SQYSKVLNLTQLKQLVDAQKENSVSITEHLQVITTLTAKEMEKEKISNLEKHLASKEV 772
QY 298 KIGADTKTEPLTRSFNEQAKGLGNDGKGRNYYDFLKGVOQOVENACNLKLTYPGLDMK 357
Db 773 EVAKLEKQLL-----BEKAAMTDAMVPRSSYEKLOQSSLESESVSLAS-KLKESYKKEK 825
QY 358 FNSLI--MLSEEQKVERDIKLE--LYSK-----YNQLOIQSSSFDNEELAILK 401
Db 826 VHSEVQIRSEVSQVREKENIQTLLKSKQEVNELLQKFOQAQBEAEMK 876

RESULT 8
US-11-052-554A-121
; Sequence 121, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR FILING DATE: 2005-02-07
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 121
; LENGTH: 1167
; TYPE: PRT
; ORGANISM: Helicobacter pylori J99
US-11-052-554A-121

Query Match 5.4%; Score 112; DB 7; Length 1167;
Best Local Similarity 21.6%; Pred. No. 1.7; Mismatches 165; Indels 114; Gaps 19;
Matches 95; Conservative 65;
QY 30 IENAECDCKNGLILESAAHNVSEVELARLKVQLKN---ALLNCIIISYRFHGIGVY--LVK 84
Db 626 ERLEKDVAKN---LESKSGNKNKMEAKSQANSQKDEIFALINKEANRARAIAAYANLKG 682
QY 85 TKOTLIDLEQPVNTELPIGFEYLDY--VROLGVDFDHIT 123
Db 683 IKRELSDKLENINKDLKDFSKSFD-EFKNGKNKDFSKAEETLKALKGSKVKGIDGINPEWIS 741
QY 124 YKVKSNKNNSLDVAKIHKSLIYENFDYILKRYVPCYTESFLLDIYLFYKIVYIEIRI 183
Db 742 ---KVENLNAALNEFFKNGK-----NKDFSKVTQAKSDLENSIKDVIINOK----ITDK 787
QY 184 IENHNF-----FYKDESLVOLQDASSATTSLSALTOS--NND 220
Db 788 VDNLNQAVSVAKATGDFSGVEQALADLKNFSKEQLAQQAQKNEDPNTGKNSALYQVNG 847
QY 221 RGSGLSFLRKQNSNNHSHKDISNLRNNDLSQELARLKNLNNEGMPFYTATPSASLEV 280
Db 848 VNGTLVGNGLSKABATTLTKNFSDDIKK---ELNAKLGNFNNN--NNNGLENSTEP----- 897

Qy 281 IKYDLSYLKEALAIKAKIGADTKEPLTRSFNEQAKGLGNDGKGRSNYYDFLKGVQBOV 340
Db 898 -----IYTQVAKKYGKID-----LDQIATASGLGVQQA-----ASFLLKRDHKV 937
Qy 341 ENACNLKLT-----YFGLDMKFNSLIMLSEBQKVE-----RDKLIELYSKYNQLI 387
Db 938 DDLSKVLGSANHEPIYATID-DLGGPPFLKRDHKVDLDSKVLGSRQKLTQKIDNINQAV 996
Qy 388 ---QSSSFDNBEALIKK 403
Db 997 SEAKASHFDNLDQMDIKL 1015

RESULT 9

US-11-155-288-20
; Sequence 20, Application US/11155288
; Publication No. US20060008468A1
; GENERAL INFORMATION:
; APPLICANT: Chiang, Chih-Sheng
; APPLICANT: Simard, John J.L.
; TITLE OF INVENTION: COMBINATIONS OF TUMOR-ASSOCIATED
; ; TITLE OF INVENTION: ANTIGENS IN DIAGNOSTICS FOR VARIOUS TYPES OF CANCERS
; FILE REFERENCE: MANUK 050A
; CURRENT APPLICATION NUMBER: US/11/155,288
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: 60/580,969
; PRIOR FILING DATE: 2004-06-17
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 976
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-155-288-20

Query Match 5.4%; Score 111; DB 7; Length 976;

Best Local Similarity 23.3%; Pred. No. 1.6;
Matches 102; Conservative 52; Mismatches 161; Indels 122; Gaps 19;

Qy 12 KISSLELYKYSIFFRNYIENVAEDCLKGLILESAAHNVSEVELARKVL--KVALINC 69
Db 406 KILTMLOKKS-----SELEMTK-----LTNNKEVEELKVLGKERTLL-- 447
Qy 70 IISYRFHGIGYVVKTDLDLEOPVNIELPIGFEYLDYEVVDLGVDFDHT--Y 124
Db 448 YENKOFEKIAELKGTQELIGLLOAREKE-----VHDLQLTAITTSQYI 495
Qy 125 KYKSNKNNSLDVAKIHSRLIYENFDYILKRYVPCYTESFLL-----DIYLFKIIY 177
Db 496 SKEVKDLKTELENEKLNKTELTSCHCNKLSLENKELTQETSDMTLEKKNQOEDINNKKQE 555
Qy 178 VEIERRIEN-----HNFLPYKDESIVQLQDALSSATTSLSALTQNNDRGSGILSSP 229
Db 556 ERMUKQIENLOETQLRNELEYVREELKQKRDVK-----CKLDKS----- 597
Qy 230 LKQNSNHSKDINSRLNDSLSQELARKLSNANNEGMFYTATPSASLEVIKYDLSVLK 289
Db 598 --EENCNLRKQVENKNKYIEELOQENKALKKKGTAE-----SKQLNVEYIKVKNLE 647
Qy 290 EALALIKAKIG--ADT--KEPLTRGFNEQ-----AKGLNGDKGRSNYYDFLKGVQ 337
Db 648 LELESAKQKFGBITTYQKEIEDKKISEENLLEVEKAKVIADAE-----VKIQ 696
Qy 338 EOVENACNLKTYFGDMKFNSLIMLSEBQKVERDIKILIE-----LYSKYNLIQSS 390
Db 697 KEIDKRCQHIAE-----MVALMEKHQYD-KIIEERDSELGLYKSEQSSSL 745
Qy 391 SFDNE-ELAILKEKLF 406
Db 746 RASLELSNLKAEELS 762

RESULT 10

US-11-077-550-112
; Sequence 112, Application US/11077550
; Publication No. US20050244435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 05/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 112
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-11-077-550-112

Query Match 5.2%; Score 107.5; DB 7; Length 888;

Best Local Similarity 20.4%; Pred. No. 2.5;
Matches 99; Conservative 65; Mismatches 146; Indels 175; Gaps 22;

Qy 6 KTKUIDKISSLELYKYSIFFRNYIENVAEDCLKGLILESAAHNVSEVELAR----- 57
Db 474 KNSFDDLSKNERIYNT-QSNYIEN---DPPINELILDIDL--ISKIELPRENTESLTD 527
Qy 58 -----LKVLKVALINCILISYRFH 76
Db 528 FNDVPVVEKOPAIKKIIFTDENTIFQYLYSQTFPFDLDIRDISLTSFDDALLFSNKVYGF 587
Qy 77 GIGYVVKTKTDLDLEQPVNIELPIGFEYLDYEVVDLGVDFDHTYKVKSNKNNSLD 136
Db 588 SMDY--IKTANKVVEAG-----LFAGVVKQIVNDF-----VIEANKSNTMD 626
Qy 137 AVKIHKSRLLI-----YEN-FDY-----ILKRYVP-----CYTESFLLDI 170
Db 627 --KIADISLIVPYIGLALNVGNETAAGNFENAFIAGASILLEFIPELLIPVVGAFILLES 684
Qy 171 YL--FEKIYVEIERRIENHFLFYKDESIVQLQDALSSATTSLSL-----SALTQSNDRSG 224
Db 695 YIDNKNKIITIDNALTKRNEKSDMWGLIYVQ--WLSTVNTQFYTIKEGMYKALNYQQA 743
Qy 225 1-----LSSFLRKQNSNHSKDINSRLNDSLSQELARKLSNANNEGMFY---TA 272
Db 744 LEEIKRYNYIYSEKSNINIDFNDS--KLNEGIAQIDNINNFINGCSVILMKM 801
Qy 273 TPSASLEVIKYDLSYLKEALAIKAKIGADTKEPLTRSFNEQAKGLGNDGKGRSNYYD- 331
Db 802 IPLAVEKLLDFDNLKKNLL-----NYIDE 826
Qy 332 ---FLKGVQBOVENACNLKTYFGDMKFNSLIMLSEBQKVERDIKILIELYSKYNLIQ 388
Db 827 NKLYLGSAEYKES-----KVNKYLKTIIMPFDLSIYTN-----DTILIEFNKYNSEIL 875
Qy 389 SSSFD 393


```

Db 78 LEDTKLTLODDTTEKIELFK-KLEFKQLLADIDTSSTN-----EEVIDKTEIEQDFQNV 131
Qy 61 QLKALLNCIIISYRPHGIGYVLVTKDTLIDLEQPVNIELPIGFEYLDYEVV----RDLG 116
Db 132 DL-NDLNEAVIHFELEGNYL-----KDTI-----LKPGF-YTNHQHVINAEDV- 174
Qy 117 VDFDHITYKVKSNKNKNSL-DAVKIHS--RL-IYEN--FDYILKRY----- 158
Db 175 KDYKHLVQWLEDKNTTKIVYDAKTYVSAHRLGINIENIEFDVMLASYIIDPSRIDDVK 234
Qy 159 --VPCYTESFLD-IYLF---EKIYVEIERRIENHFLFYKDESLVQLQDALSSATTSL 212
Db 235 SVVSLYGQNYVKDNITIFGKGKHHIPPEPIINEH-----IASVTEAIAAVPTMK 285
Qy 213 ALTQSNDRGSGI-----LSSFLRKQNSNHSKDISNLRNLNDSLSQELAKKSNL 263
Db 286 SOLEDYNO-----IELLKDELPLARILSEMEIEGIYTDINDLKEMEFIOKKLDVLI 341
Qy 264 N-NEGMFYTATPSASLEVIKYDLSYLKEALALIKAKIGADT-----KEPL---TR 309
Db 342 HESAGEAFNINSPKQGVVLPETLQLP---VIKKTGYSTAVDVLEKLOGEHPIIDIL 398
Qy 310 SFNEQAKGLGNDKGDRSNYDFLKV---QEQVENACNLKT---KYFGMDMKFNSL-I 362
Db 399 EYROLAK-----LQSTYVEGLOKVISKDHRIHTRFNQTLAQTRLSSIDPNLQNIPI 450
Qy 363 MLSEEQKVERDIKLIELYSKNQLIQSSSFONEELAIL 400
Db 451 RLEGRKIRKAFK----PTSKDSVILSADYSQIELRVL 484

```

Search completed: January 24, 2006, 20:55:33
Job time : 7.51393 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:28:54 ; Search time 15.4404 Seconds
(without alignments)
2536.214 Million cell updates/sec

Title: US-10-688-058-4

Perfect score: 2062

Sequence: 1 MCDLRKTKLIDKISSELYK.....QSSSFDNEELAILKEKLPSP 407

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1019.5	49.4	413	2 F70211	hypothetical prote
2	186.5	9.0	3394	2 T18501	hypothetical prote
3	150.5	7.3	1957	2 T38077	hypothetical coile
4	148	7.2	2269	2 T38677	rhoetry protein -
5	146	7.1	1827	2 A35694	cuti protein - fis
6	146	7.1	1828	2 T41455	conserved hypothe
7	145.5	7.1	971	2 C82880	hypothetical prote
8	144.5	7.0	2166	2 G70163	hypothetical prote
9	143.5	7.0	1169	2 A64505	p115 homolog - Met
10	141	6.8	819	2 E70105	p115 protein homol
11	140.5	6.8	864	2 B90395	purine Ntase [imp
12	140	6.8	703	2 T64226	DNA helicase II (m
13	137.5	6.7	1127	2 T28317	ORF MSV156 hypothe
14	137.5	6.7	1979	2 C71622	hypothetical prote
15	136	6.6	3724	2 T18427	hypothetical prote
16	135.5	6.6	706	2 D90124	hypothetical prote
17	134	6.5	1939	2 T18372	repeat organellar
18	133.5	6.5	2663	1 S28261	centromere protein
19	133	6.5	546	2 A99600	conserved hypothe
20	132.5	6.4	1780	2 T17272	hypothetical prote
21	131.5	6.4	1738	2 T14867	interaptin - slime
22	130.5	6.3	1073	2 S14032	kinesin-related pr
23	130.5	6.3	1085	2 T38378	kinesin-like prote
24	129.5	6.3	365	2 JC2559	flagellin flic-1 -
25	129.5	6.3	488	2 AD0735	conserved hypothe
26	129	6.3	2910	2 C89910	DNA-directed RNA p
27	128.5	6.2	1009	2 C89910	hypothetical prote
28	128	6.2	1183	2 F90559	conserved hypothe
29	128	6.2	2010	2 B71616	phosphatase (acid

30	127.5	6.2	488	2 AG0621	probable bacteriop
31	127.5	6.2	689	2 T18494	hypothetical prote
32	126.5	6.1	2401	2 T28676	rhoetry protein -
33	126	6.1	611	2 D82881	zinc metalloprotei
34	126	6.1	980	2 E71606	hypothetical prote
35	126	6.1	1138	2 A82939	membrane nuclease
36	125	6.1	1365	2 T18419	hypothetical prote
37	125	6.1	1875	2 S38173	myosin-like protei
38	125	6.1	2819	2 T09080	probable chloroqui
39	124.5	6.0	944	2 S26710	spindle pole body
40	124.5	6.0	1021	2 E64576	hypothetical prote
41	124.5	6.0	1272	2 H82926	conserved hypothe
42	124.5	6.0	1294	2 T18473	hypothetical prote
43	124	6.0	1031	2 H81288	probable sugar tra
44	123.5	6.0	871	2 E97035	DNA polymerase I,
45	123	6.0	772	2 B82888	phenylalanine-tRNA

ALIGNMENTS

RESULT 1

F70211

hypothetical protein BBA38 - Lyme disease spirochete plasmid A/lp54

C/Species: Borrelia burgdorferi (Lyme disease spirochete)

C/Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004

C/Accession: F70211

R/Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White

son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,

i Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A/Authors: Smith, H.O.; Venter, J.C.

A/Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A/Reference number: A70100; MUID:98065943; PMID:9403685

A/Accession: F70211

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-413 <KLE>

A/Cross-references: UNIPROT:O50931; UNIPARC:UPI000005691F; GB:AE000790; NID:g2690224; PI

A/Experimental source: strain B31

C/Genetics:

A/Genome: plasmid

C/Superfamily: Borrelia burgdorferi hypothetical protein BBA38

Query Match	49.4%	Score	1019.5	DB 2	Length	413			
Best Local Similarity	53.0%	Pred. No.	1.2e-52						
Matches	212	Conservative	72	Mismatches	99	Indels	17	Gaps	7
QY	12	KISSELYKYSIFPRNYIENVAEDCLXNGL---	ILESAAHVSEVELARLVKQKVALIN	68					
Db	20	KINPIDVYRYSIFPRNYIESTAEDALNGINLSLLEKSSKSLGLKVLKEALLE	79						
QY	69	CIISVRFHGIGYLVKTKTLDLLEQPNVIELPGFVLDYVYRDVGDFDHYTKVKS	128						
Db	80	AMISYRFNGAGYILVLPKASEDELSKVNSELPTGFKYLDFOKINKR-DSSVVEY--LS	136						
QY	129	NNKQ-----NSLDAVKIHKSLRIIYENFDYILKRVPCYTESFLDIYLFEBKIYVEIERI	184						
Db	137	NSKPDPPERARVVKIDKSRVIIYENYDVLGEQEPAYQTSSLNLCLEQIYLEIERI	196						
QY	185	ENHNFYKDSRLVQLQDALSSATTSLSALTQSNDRSGSLSSFLRKQNNHNSKDISN	244						
Db	197	RNYNFLFYKDHVLGLVESLEIAKEINVLA---NSKQK-IFSTFFKQAEFN---KSFQA	249						
QY	245	LRNLDISIQLARLKNLNNEGFFYATPSASLEVIKYDLSYLKEALAIKAKIGADTK	304						
Db	250	LSSVDSLELSKIKNTLNNDGIFYTASENARLEVIKYDLEFLKDAFELVAKIGADTK	309						
QY	305	EPLTRSFNEQAKGLGNDGKDRSNYYDFLKGQVQEVENACNLKTKYFGLDMKFNLSIML	364						
Db	310	EPLTRSFNEQVKGIGSSGKGKSNYYDYLVKGQESVANACNLKLNKYRLNKNFNEALEAL	369						
QY	.365	SEQKVERDIKILIELYSKYNQLIOSSSPDNNEELAILKEKL	404						

Db 269 IDVEIENL--KURLNNIINELNEKNEVEVLELHKS-----IKLEVEIENDKKVL 316
QY 169 D---IYLFKIVYVEIERR-----IENHFLFYKDESIVQ-----LQ 201
Db 317 DSSINELKKVEIEENKKKEIKETQKIIENRDSIIKEEQIKEEIKKLNLYEKERLK 376
QY 202 DALSSATTSLSALTOSNDRGSGILSSFLRKONS-----NNHKSIDSINLND----- 250
Db 377 EIAESESIIKHLKESEME---IADETAKQNELYRLKCELDNDLNIRKNKFNIEKVN 432
QY 251 ----SLSOELARLK-----SNLNNEGMFTYATPSASLEVIKYDL-----S 286
Db 433 EMIKKLEKELEVEDVTKPYLELENLNV---IEFSKGIKELEKKELQAKLDELHAE 491
QY 287 YLKEALALIKA-----KIGADTKPLTRSFNPOAKGL-----GNDGK-----G 324
Db 492 YVKN-ARIKALKEMEELSMD--RAIREITLNNALPGIIDIVGNLTKTIEYKTAJEVAAG 548
QY 325 DRSNYDFLKGVOQOVENACNLKLTG-----YFGILD-MKFNSLIMLSEQKVERDIKLE 378
Db 549 NRELNI-VVGRMDDAVRAIKYKELKGRATPLPLDRIEGREAYVIDEGVIGRAIDLVE 607
QY 379 LYSKYNQLOIQQSSFFN-----BELAILKE 402
Db 608 FDEKYRRVFE-YVGNVVVENIDIAKE 634

RESULT 10
E70105
P115 protein homolog - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C:Accession: E70105
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete.
A:Reference number: A70100; PMID:98065943; PMID:9403685
A:Accession: E70105
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-819 <KLE>
A:Cross-references: UNIPROT:O51074; UNIPARC:UPI000005731A; GB:AE001118; GB:AE000783; NID
A:Experimental source: strain B31

Query Match 6.8%; Score 141; DB 2; Length 819;
Best Local Similarity 22.2%; Pred. No. 0.84;
Matches 117; Conservative 93; Mismatches 152; Indels 164; Gaps 30;

QY 4 LK-KTKLIDKISSLE--LYKYSIFPRNYIENVAEDCLKNGILSSAAHNVS---EVELA 56
Db 216 LRKHQDUSQKLEALEKTITLKKLFNINFDLNV---LKSENLGLSLFSLSSGPKERLE 271
QY 57 RLKVLQKALLNCIIISYRPHGIGYVLVKTOTLIDLEQPVNIETLPGFYLDYEVYRDLG 116
Db 272 FYSFREKVINN-----IQELEKDLMLKSKILGLEIKIQKLEQKTLKN 317
QY 117 VPDHTYTKVKNKNKNSLDVAKIHKSLRIIYENFDYI-LKRYVPCTYSP--LDDIYLF 173
Db 318 LEDKPKVKSSESEKKNNSNNLYQLNLSLKEKNDFFSLSDINRYTKSPFELVDLIL- 376
QY 174 EKIYVEIERIENHNFYKVD---ESLVLQDALS-----SATTLSALT 215
Db 377 -----SVLRSAKSEEFVLKHEHLDLSLKFEESSLSIKYIKIRVNLKIYKDNELALL 431
QY 216 QSNDRGSGI-----LSSFLRKONS--NNHKSIDSINLR-----NLNDSLSO----E 255
Db 432 R---DKIEPIFDQVNLKFLLEKNAKSNLAKNEITNIERLEETKTLQNDVLGELEYIE 488
QY 256 LARLKSNI-----NNEGMFTYATPSASLEVIKYDLSYLK-EALALIKAKIGADTKEP 306

RESULT 12

Db 489 LSKFK-NLDEIKLIDGNLEFLFESKN-----SLDEELKOL-YLKLENLNLKESDIQLNLNL 543
QY 307 L-----TRSFNEQ---AKGLNDGKGRSNRYDFLK-----GVQOVEN-- 342
Db 544 IGASKFSSESPKXDFNSLVPLNDFK-KTNYEYIKQTEYEFLLNSVNSIKSEIENFN 601
QY 343 -----ACNLKLTQYFGLDMKFNSL-----IMLSEE--- 367
Db 602 VSNEMSKFEELEDLATRELLRKEIDAIKLGDYVFFNIDKEFDTEKFEKVSQVLEDLK 661
QY 368 -----OKVERDIKILIELYSKYNQLOIQQSSFFN--NEELAILKEKLF 405
Db 662 LSKYSLQKLOKKIK-NEIYKFKFE-----SFDEINKNFSFFFKLIF 701

RESULT 11
B90395
purine NTPase [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: B90395
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: B90395
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-864 <KUR>
A:Cross-references: UNIPROT:Q97WH0; UNIPARC:UPI000013302A; GB:AE006641; NID:gl3815551;
C:Genetics:
A:Gene: SSO2249

Query Match 6.8%; Score 140.5; DB 2; Length 864;
Best Local Similarity 20.4%; Pred. No. 0.96;
Matches 87; Conservative 83; Mismatches 150; Indels 107; Gaps 19;

QY 6 KTKLIDKISSLELYKYSIFPRNYIENVAE-----DCLKNGILSSAAHNVSSEVLARLKV 60
Db 206 KEKLEDEIKNLE-----KRIKIDQDFEYKKNQYIKLTTILKIGELNELNR 256
QY 61 QLKALLNCIIISYRPHGIGYVLVKTOTLIDLEQPVNIETLPGFYLDYEVYRDLGVDPD 120
Db 257 SIEE-----LRQTNEMDQLEKEIN-----ELENLRLNKLKPE 289
QY 121 HITYKVNKNKNSLDVAKIHKSLRIIYENFDYILKR---YVPCYTESFLDDIYLFKIY 177
Db 290 KYEVLAKSHTE-MSANVINLEK-----IEYEKAIRKEELEPEKYLK-----Y 332
QY 178 VEIERRIENHNFYKDESILVQLDALSATTLSALTOSNDRGSGILSSFLRKONSNN 237
Db 333 KELEKLESE---LQPKYQYVLEKLDLDSKLNKLERL-----EKDASELSNDIDKVNLSL 384
QY 238 HSKDISINLNDLSLSOELARLKNLNNEGMFTYATPSASLE-----VI 281
Db 385 QKVETRRKKQLN--LRAQLAKVESLISEKNEIINNIQSVEGETCPVCGRPLDEEHKQKII 442
QY 282 KYDSLKEALALIKAKIGADTKPLTRSFNE---QAQGLNDGKGRSNRYDFLKGQVE 338
Db 443 KEAKSYIIQ-LLELNKELEEBLKK-ITNELNKIEREYRLSN--KASVDNVRQLKCLNE 499
QY 339 QVENACNLKLTQYFGLDMKFNSLIMLSEE-OKVERDIKILIELYSKYNQLOIQQSSSDNEEL 397
Db 500 EIE-----LHSEIBSLKNIDBEIKKINEVEKELKUY---YEEFMRLSKYTKKEEL 546
QY 398 AILKEKL 404
Db 547 DKERVKL 553

I64226

DNA helicase II (mutB1) homolog - Mycoplasma genitalium
 C:Species: Mycoplasma genitalium
 C>Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 09-Jul-2004
 C:Accession: I64226; S18697
 R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;
 M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.
 , C.A.; Venter, J.C.
 Science 270, 397-403, 1995
 A:Title: The minimal gene complement of Mycoplasma genitalium.
 A:Reference number: A64200; MUID:96026346; PMID:7569993
 A:Accession: I64226
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-703 <TIG>
 A:Cross-references: UNIPROT:P47486; UNIPARC:UPI0000137EPD; GB:U39703; GB:L43967; NID:938
 A:Experimental source: strain G-37
 R:Peterson, S.N.; Schramm, N.; Hu, P.; Bott, K.F.; Hutchison III, C.A.
 Nucleic Acids Res. 19, 6027-6031, 1991
 A:Title: A random sequencing approach for placing markers on the physical map of Mycopla
 A:Reference number: S18693; MUID:92051396; PMID:1945886
 A:Accession: S18697
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 277-345 <PET>
 A:Cross-references: UNIPARC:UPI000016FA46; EMBL:X61517; NID:944318; PIDN:CAA43729.1; PID
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
 C:Genetics:
 A:Genetic code: SG3
 C:Superfamily: helicase II

```

Query Match          6.8%; Score 140; DB 2; Length 703;
Best Local Similarity 21.3%; Pred. No. 0.79;
Matches 105; Conservative 73; Mismatches 140; Indels 176; Gaps 26;

QY 4 LRKTKLIDKISSLELYKYSIFRNVIENVAEDCLKNGLILESAAHNVSEVELARLKVQ-- 61
Db 94 LKRRFIID-----DISYFLKEFLADSKLDINLAKQIDFNKTFADFEINKLDQDER 146
QY 62 -----LKNALLNCIISYR-----FHGIGYVLVKT----- 85
Db 147 LISLCEHSLNKKDEYSTLKTQLINAFISYERKNKILNNKLDPHDL---LIKTCLNLSNDN 203
QY 86 -----KOTLIDLEQVPNIELPIGPEYLDYEVVRDLGVDFDHIYVKVKNK----- 131
Db 204 DLLNQWSEQFQHLVDFQDTY-----QIQYELIKML-VTKKNFLVGDNNQMIYR 254
QY 132 -----NNSLDVAK-----IHKSRLIITYENPDYILKRYVPCYTESFLDIYLPFK----- 175
Db 255 MRGAVNGIITALKHDFNVFKSN-EFFINQYRCDQNILAVANQILKIMAYEKQVKTEN 313
QY 176 -----IYVEIERIENH-NFLPYKDESILVOLQDALSATTSLTSALTQSNDRG 222
Db 314 LLFSTLNSDKKPYVQAE-SVENQANWIFNKIKALNQ-----TEKINFKD 357
QY 223 SGLSFLRKQNSNNHSDKISLRLNLSLQ-ELARLKSNNLNEGMYFATPSASLEVI 281
Db 358 MAIL---FRK-----NRDITTWELIEADGTIPLPKQSYFN-----QLV 394
QY 282 KYDSLVLKEALALIKAKIGADTKETPLTRSFNBQAKGLNDGK-----GDRSNYYDFLKG 336
Db 395 K-----LQRLVIAISRTNLDIKRAL-----QALKWSNDLKLWQSKTNLDFLWKS 444
QY 337 Q-EQVENACNLKLYKFGLDKFNLSLMLSEBQKVERDIKLIELYS--KYNQLIOSSSPD 393
Db 445 ELNQKNHSSKLKATGY-----FNLLIKLAEDQOI--NLLFTFLPKLKVQDQTIEN---- 492
QY 394 NEELAILKELKFSF 407
Db 493 -----LLWKKLTLEF 501

```

RESULT 13

T28317

ORF MSV156 hypothetical protein - Melanoplus sanguinipes entomopoxvirus
 C:Species: Melanoplus sanguinipes entomopoxvirus
 C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
 C:Accession: T28317
 R:Afonso, C.L.; Tullman, B.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.
 J. Virol. 73, 533-552, 1999
 A:Title: The genome of Melanoplus sanguinipes entomopoxvirus.
 A:Reference number: 220484; MUID:99102612; PMID:9847359
 A:Accession: T28317
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1127 <AFO>
 A:Cross-references: UNIPROT:Q9YVT6; UNIPARC:UPI00000F6900; EMBL:AF063866; NID:94049647;
 C:Genetics:
 A:Note: MSV156

```

Query Match          6.7%; Score 137.5; DB 2; Length 1127;
Best Local Similarity 20.6%; Pred. No. 2;
Matches 104; Conservative 86; Mismatches 168; Indels 147; Gaps 26;

QY 8 KLIDKISSLELYK-----YSIFRNVIENVAEDCLKNGLILESAAHNVSEVELAR 57
Db 407 KKIDDIKNNLQKLEESYKKIDQEYTKYKKNKEYND-----IIEKNNLQKLEEN 460
QY 58 LKVQ-----LKNAL---LNCIISYRFHGIGYVLVKT---DTLIDLEQPV--NIELPIGF 104
Db 461 KKIDQEYTKYKKNKEYNDIIEKNNLQKLEENKKNINDKLTCLKNDIESNTEL---P 517
QY 105 EYLDYEVYRD-----LGVDFDHITYKVKSN-NKNNSLDAVAKIHKSLI-----IYEN-- 150
Db 518 NKLNISDFKDKSREIAKLNTEYEQLRKDLLENINKTNEL--MKLSDNKLSSLEQLYDSKK 575
QY 151 --PDYILKRYVPCYTESFLDIYL--PEK--IYVEIER--IENHNFIFY 192
Db 576 NILDGIDKIYNSLKEKNDKIDYFENTEKFDIYVNIENKFGNLDISIINKINNDQPKBY 635
QY 193 -----KDESILVOLQALSATTSLSALTQSNDRGSGILSSFLRKQNSNNHSDKISNL 245
Db 636 INSKIDSKNSELSTWFDIFNAKQIASIT--NN-----IENSKINKIDNEF 681
QY 246 RNLDLSLSQELA-----RLKSNLNNEGMYFATPSASLEVIKYDLSYLSKEALAL 294
Db 682 IISNEDSSKELLDIRKYKQKQDKIKDAMNTEVKSFENTLQKIDISIKSNINELNAYDI 741
QY 295 IKAKIGADTKETPLTRSFNBQAKGLNDGKGRSNYYDFLKGQVEVACNLKTYFGL 354
Db 742 INTR-----ANDLDDKLNNYGEFKNLYNNASDLDDTIQKNDEKVK-QLNEYLEK 791
QY 355 D-----MKFNSLIMLSEBQKVERDI-----KLIELYKYNQL-- 386
Db 792 NKQNSIENDIVNFIKELIKFN--TETNKSINELLTNDINDKIFKLYKELNKIST 847
QY 387 -----IQSSSPD--NEELAILKEKL 404
Db 848 NLLKIYKNEIDNVNKLKLSIVIENL 872

```

RESULT 14

hypothetical protein PFB0145c - malaria parasite (Plasmodium falciparum)
 C:Species: Plasmodium falciparum
 C>Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004
 C:Accession: C71622
 R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
 ; Partea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
 Science 282, 1126-1132, 1998
 A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
 A:Reference number: A71600; MUID:99021743; PMID:9804551
 A:Accession: C71622
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1979 <GAR>

This Page Blank (uspto)


```
Matches 400; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY 1 MCDLRKTKLIDKISSLELYKYSIFPRNYIENVAEDCLKNGLILESAAHNVSEVELARLKV 60
DB 1 MCDLRKTKLIDKISSLELYKYSIFPRNYIENVAEDCLKNGLILESAAHNVSEVELARLKV 60
QY 61 QKQALLNCIIISYRPHGIGYLVKTKDTLIDLEQPNVIELPIGFEYLDYVYRDIGVDFD 120
DB 61 QKQALLNCIIISYRPHGIGYLVKTKDTLIDLEQPNVIELPIGFEYLDYVYRDIGVDFD 120
QY 121 HITKVKSNKNNKNSLDVAKIHKSRLLIYENFDYILKRYVPCYTESFLDIIYLFKIIYVEI 180
DB 121 HITKVKSNKNNKNSLDVAKIHKSRLLIYENFDYILKRYVPCYTESFLDIIYLFKIIYVEI 180
QY 181 ERIENHNFLFYKDESIVQLDALSSATTSLSALTQGNNDRGSGILSSFLRKQNSNNHSHK 240
DB 181 ERIENHNFLFYKDESIVQLDALSSATTSLSALTQGNNDRGSGILSSFLRKQNSNNHSHK 240
QY 241 DISNLRNLSLSQELARLKSNNLNNEGMYTATPSASLEVIKYDLSYLKEALALIKAKIG 300
DB 241 DISNLRNLSLSQELARLKSNNLNNEGMYTATPSASLEVIKYDLSYLKEALALIKAKIG 300
QY 301 ADTKEPLTRSFEQAKGLGNDGKDRSNYYDFLKGVOEQVENACNLKTYFGLDKMFNS 360
DB 301 ADTKEPLTRSFEQAKGLGNDGKDRSNYYDFLKGVOEQVENACNLKTYFGLDKMFNS 360
QY 361 LIMLSEEQKVERDIKLIELYSKYNQLIQSSSFDNEELAILKEKLFSP 407
DB 361 LIMLSEEQKVERDIKLIELYSKYNQLIQSSSFDNEELAILKEKLFSP 407

RESULT 4
Q9S0A2 BORBU PRELIMINARY; PRT; 407 AA.
AC Q9S0A2;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE Hypothetical protein BBO01.
GN OrderedLocusNames=BBO01;
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=139;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=20138354; PubMed=10672174;
RA Casjens S., Palmer N., van Vugt R., Huang W.M., Stevenson B., Rosa P.,
RA Lathigra R., Sutton G.G., Peterson J.D., Dodson R.J., Haft D.H.,
RA Hickey E.K., Gwinn M.L., White O., Fraser C.M.;
RT "A bacterial genome in flux: the twelve linear and nine circular
RT extrachromosomal DNAs in an infectious isolate of the Lyme disease
RT spirochete Borrelia burgdorferi."
RL Mol. Microbiol. 35:490-516(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A.,
RA Artach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi."
RL Nature 390:580-586(1997).
DR EMBL; AF001579; AAF07602.1; -; Genomic_DNA.
DR TIGR; BBO01; -;
```

```
DR InterPro; IPR009427; DUF1073.
DR Pfam; PF06381; DUF1073; 1.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 407 AA; 47061 MW; 8E34A6BF392D15EE CRC64;

Query Match
Best Local Similarity 97.9%; Score 2019; DB 2; Length 407;
Matches 397; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
QY 1 MCDLRKTKLIDKISSLELYKYSIFPRNYIENVAEDCLKNGLILESAAHNVSEVELARLKV 60
DB 1 MCDLRKTKLIDKISSLELYKYSIFPRNYIENVAEDCLKNGLILESAAHNVSEVELARLKV 60
QY 61 QKQALLNCIIISYRPHGIGYLVKTKDTLIDLEQPNVIELPIGFEYLDYVYRDIGVDFD 120
DB 61 QKQALLNCIIISYRPHGIGYLVKTKDTLIDLEQPNVIELPIGFEYLDYVYRDIGVDFD 120
QY 121 HITKVKSNKNNKNSLDVAKIHKSRLLIYENFDYILKRYVPCYTESFLDIIYLFKIIYVEI 180
DB 121 HITKVKSNKNNKNSLDVAKIHKSRLLIYENFDYILKRYVPCYTESFLDIIYLFKIIYVEI 180
QY 181 ERIENHNFLFYKDESIVQLDALSSATTSLSALTQGNNDRGSGILSSFLRKQNSNNHSHK 240
DB 181 ERIENHNFLFYKDESIVQLDALSSATTSLSALTQGNNDRGSGILSSFLRKQNSNNHSHK 240
QY 241 DISNLRNLSLSQELARLKSNNLNNEGMYTATPSASLEVIKYDLSYLKEALALIKAKIG 300
DB 241 DISNLRNLSLSQELARLKSNNLNNEGMYTATPSASLEVIKYDLSYLKEALALIKAKIG 300
QY 301 ADTKEPLTRSFEQAKGLGNDGKDRSNYYDFLKGVOEQVENACNLKTYFGLDKMFNS 360
DB 301 ADTKEPLTRSFEQAKGLGNDGKDRSNYYDFLKGVOEQVENACNLKTYFGLDKMFNS 360
QY 361 LIMLSEEQKVERDIKLIELYSKYNQLIQSSSFDNEELAILKEKLFSP 407
DB 361 LIMLSEEQKVERDIKLIELYSKYNQLIQSSSFDNEELAILKEKLFSP 407

RESULT 5
Q9S057 BORBU PRELIMINARY; PRT; 408 AA.
AC Q9S057;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE Hypothetical protein BBN01.
GN OrderedLocusNames=BBN01;
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=139;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=20138354; PubMed=10672174;
RA Casjens S., Palmer N., van Vugt R., Huang W.M., Stevenson B., Rosa P.,
RA Lathigra R., Sutton G.G., Peterson J.D., Dodson R.J., Haft D.H.,
RA Hickey E.K., Gwinn M.L., White O., Fraser C.M.;
RT "A bacterial genome in flux: the twelve linear and nine circular
RT extrachromosomal DNAs in an infectious isolate of the Lyme disease
RT spirochete Borrelia burgdorferi."
RL Mol. Microbiol. 35:490-516(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A.,
```



```
Db 1 MCNLRKTKLVDRINLLKLYRYSIFRNVIENVAEDCLNGLVLESTSHNVSEFELDRKV 60
QY 61 QLNKALLNCIIISYRFH--GIGVVLVKTOTLDLEQPVNIELPIGFYLDYEVYRDGLVD 118
Db 61 QLNKALLNCIIISYRFHNGIGVVLVKTQDALDLEPNIELPIGFYLDYEVSKDSGID 120
QY 119 FPHITVYKVNKNNSDLDAVHKHRLIYENFDYILKRYVPCYTESFLDIIYLPKIIYV 178
Db 121 FDIYINVKPNKNGSNFNTVKHKHRLIYENFDYILKRYVPCYTESFLDIIYLPKIIYV 180
QY 179 ETERIENHNFYKDESIVQDALSSATTSLSALTO--SNNDRSGILSSFLRKQNSNN 237
Db 181 ETERIENHNFYKDESIVQDALSSATTSLSALTO--SNNDRSGILSSFLRKQNSNN 240
QY 238 HSKDLSNRLNDSLSQELARLKNLNNEGMFTATPSASLEVIKYDLSYLKEALALIKA 297
Db 241 HSKDLSNRLNDSLSLAHELLRLKNLNNEGMFTATPSASLEVIKYDLSYLKEALALIKA 300
QY 298 KIGADTKPELTSFNEQAKGLNDGKDRSNYYDFLKGVOE 338
Db 301 KIGADTKPELTSFNEQAKGLNDGKDRSNYYDFLKGVOE 341

RESULT 8
Q9KKC1 BORHE
ID Q9KKC1 BORHE PRELIMINARY; PRT; 404 AA.
AC Q9KKC1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Hypothetical protein.
OS Borrelia hermsii.
OG Plasmid cp32.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=140;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HS-1;
RX MEDLINE=20316012; PubMed=10858201;
RX DOI=10.1128/JAI.68.7.3900-3908.2000;
RA Stevenson B., Porcella S.F., Oie K.L., Fitzpatrick C.A., Raffel S.J.,
RA Lubke L., Schrupp M.E., Schwan T.G.;
RT "The relapsing fever spirochete Borrelia hermsii contains multiple,
RT antigen-encoding circular plasmids that are homologous to the cp32
RT plasmids of Lyme disease spirochetes."
RL Infect. Immun. 68:3900-3908(2000).
DR EMBL; AF123078; AAF28886.1; -; Genomic_DNA.
DR InterPro; IPR009427; DUF1073.
DR Pfam; PF06381; DUF1073; 1.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 404 AA; 46409 MW; 5567119DBB024671 CRC64;

Query Match 57.0%; Score 1175; DB 2; Length 404;
Best Local Similarity 57.9%; Pred. No. 6.4e-59;
Matches 231; Conservative 69; Mismatches 89; Indels 10; Gaps 3;

QY 13 ISSLELYKYSIFRNVIENVAEDCLNGLILESAHNSEVELARLKVQLKNALNCIIIS 72
Db 7 INSRDLYKYSIFRNVIENVAEDCLNGLILESAHNSEVELARLKVQLKNALNCIIIS 66
QY 73 YRPHGIGYLVKTOTLDLEQPVNIELPIGFYLDYEVYRDGLVDYFPHITVYKVNNS-- 130
Db 67 YRPHGIGYLVKTOTLDLEQPVNIELPIGFYLDYEVYRDGLVDYFPHITVYKVNNS-- 126
QY 131 KNSLDAVHKHRLIYENFDYILKRYVPCYTESFLDIIYLPKIIYVIERIENHNL 190
Db 127 DNLSELSLHKHRLIYENFDYILKRYVPCYTESFLDIIYLPKIIYVIERIENHNL 186
QY 191 FYKDESIVQDALSSATTSLSALTO--SNNDRSGILSSFLRKQNSNNHSDINLN 247
Db 187 FYKDESIVQDALSSATTSLSALTO--SNNDRSGILSSFLRKQNSNNHSDINLN 241
QY 248 LNDLSQELARLKNLNNEGMFTATPSASLEVIKYDLSYLKEALALIKAKIGADTK 307
```

```
Db 242 ANDELMRELEKSLKNDGIFYSCEPNASLEVIKYDLSYLKEALSLVKAIGADTK 301
QY 308 TRSNEQAKGLNDGKDRSNYYDFLKGVOQVQVNAACNKLTKYFGLDMKFNLSIMLSEE 367
Db 302 TRSNEQAKGLNDGKDRSNYYDFLKGVOQVQVNAACNKLTKYFGLDMKFNLSIMLSEE 361
QY 368 QKVERDKLIELYKYNQLOIQQSSFDNEELAILKEKLF 406
Db 362 EKIRDMRLLEIYERYSSVISNPNSLDEKMKLENLFT 400

RESULT 9
O50931 BORBU PRELIMINARY; PRT; 413 AA.
AC O50931;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE Hypothetical protein BBA38.
GN OrderedLocusNames=BBA38;
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp54.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=139;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A.,
RA Artiach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi."
RT Nature 390:580-586(1997).
DR EMBL; AE000790; AAC66300.1; -; Genomic_DNA.
DR PIR; F70211; F70211.
DR TIGR; BBA38; -.
DR InterPro; IPR009427; DUF1073.
DR Pfam; PF06381; DUF1073; 1.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 413 AA; 47901 MW; 382654B2DF351241 CRC64;

Query Match 49.4%; Score 1019.5; DB 2; Length 413;
Best Local Similarity 53.0%; Pred. No. 4.5e-50;
Matches 212; Conservative 72; Mismatches 99; Indels 17; Gaps 7;

QY 12 KISSLELYKYSIFRNVIENVAEDCLNGL---ILESAAHNSEVELARLKVQLKNALIN 68
Db 20 KINPIDVTRYISIFRNVIENVAEDCLNGLINLSLEKSLKSSLSLGLVKVQLKALLE 79
QY 69 CIISYRPHGIGYLVKTOTLDLEQPVNIELPIGFYLDYEVYRDGLVDYFPHITVYKVS 128
Db 80 AMISYRPHGIGYLVKTOTLDLEQPVNIELPIGFYLDYEVYRDGLVDYFPHITVYKVS 136
QY 129 NNKN-----NSLDAVHKHRLIYENFDYILKRYVPCYTESFLDIIYLPKIIYVIERI 184
Db 137 NSKDPDFERARVVKIDSRVYIYENYDVLGEQEPAYTQTSLLNLCLEQIYLEIKRI 196
QY 185 ENHNFYKDESIVQDALSSATTSLSALTO--SNNDRSGILSSFLRKQNSNNHSDINLN 244
Db 197 RYNYNLFYKDEHVLGVLESLEAKSEINVL---NSKGK-IFSTFFKAQEPN---KSPQA 249
QY 245 LRLNDSLSQELARLKNLNNEGMFTATPSASLEVIKYDLSYLKEALALIKAKIGADTK 304
Db 250 LSSVDELSELRLSKNTLNDGIFTYTAGENARLEVIKYDLEFLDKAPFLVAKIGADTK 309
```


O77384 PLAF7
ID O77384_PLAF7 PRELIMINARY; PRT; 3394 AA.
AC O77384;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein PF0760C
GN Name=PF0760C; Synonyms=MA13P6.11;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99376085; PubMed=10448855; DOI=10.1038/22964;
RA Bowman S., Lawson D., Basham D., Brown D., Chillingworth T.,
RA Churcher C.M., Craig A., Davies R.M., Devlin K., Feltwell T.,
RA Gentles S., Gilliam R., Hamlin N., Harris D., Holroyd S., Hornsby T.,
RA Horrocks P., Jagels K., Jassal B., Kves S., McLean J., Moule S.,
RA Mungall K., Murphy L., Oliver K., Quail M.A., Rajandream M.-A.,
RA Rutter S., Skelton J., Squares R., Squares S., Sulston J.E.,
RA Whitehead S., Woodward J.R., Newbold C., Barrell B.G.;
RT "The complete nucleotide sequence of chromosome 3 of Plasmodium
falciparum.";
RL Nature 400:532-538(1999).
DR EMBL: Z98551; CAB1140.1; -; Genomic_DNA.
DR PIR: T18501; T18501.
KW Hypothetical protein.
SQ SEQUENCE 3394 AA; 402949 MW; 81B886F2056A4B57 CRC64;
Query Match 9.0%; Score 186.5; DB 2; Length 3394;
Best Local Similarity 21.5%; Pred. No. 0.13;
Matches 104; Conservative 89; Mismatches 181; Indels 109; Gaps 20;
QY 1 MCDLRTKLIDKISSLELYKYIFPRNYIENVAEDCLKNGLILESAAHNVSEVELARLV 60
Db 1561 MCDVTLASGVNKKD-----FLNLEENEENFLEIRIKSLQELCDKESILKI 1612
QY 61 Q-LKVALNCIISYRFGHIGYVLTOTLIDLEQVNIELPIGFEYL-----DYEV 111
Db 1613 KGEKNILITCIETWK-----CFCKNSKEISRLKEICKEQLEKHEFLINKSNEDKLKY 1667
QY 112 VPDLDV-----PDHITYKVSNNKNSLDV-----KHKSRLIYENFDYILKRYVPC 161
Db 1668 INSLJCDEKDYDVVVKDQIK-NMNERIDKLNNDINEKSYETKLLKHNNLNINEMNLK 1726
QY 162 YTESFLDLYLPKRYVETERR-----IEN-HNPLFVKDESLVOLQDALSATTSL 216
Db 1727 NKETENMIAKQEEYDIKIKKDKTNIQNEYNDLLEKNEVVRKNMLYN-----DMNVLLK 1783
QY 217 SNDRSGSILSFLRKQNSNNHSDISNRLN-----NDLSLQELARLKSNNNEG 267
Db 1784 EHKE-----EFLLEKENIKILQKNTVNDMFKNQINVDNNLLKNRLDQL-FNINQDL 1836
QY 268 MPTATPASPASLEVIKYDLSYLKEALAIKAKIGADTK-----EPLTRSFNQAKGL 318
Db 1837 QKHLDTNQKHLBQKYDYIEIKERLKIETKINKQEKYIIQKQDNLLNDFNTTTT 1896
QY 319 GNDGKGRDN-----YYDFLKGVOEVENACNLKTKYFGLDMKFNLSLIMS 365
Db 1897 NNNNNNDNNNDNNNDNNNDNNNTYQOFIHSKANLENS-RLELKELSN-----NEKIQUS 1951
QY 366 EQ-----KVERD-IKLEL-----YSKYNQLIOSSSPDNBELAIL 400
Db 1952 DEKNRMKITLEDKLFKNEKQKWLQIIDDNNKNVMIQNKLTNLDMLSEENMLLN 2011
QY 401 KEK 403
Db 2012 KEE 2014
RESULT 14
Q9BK46_PLAF7
ID Q9BK46_PLAF7 PRELIMINARY; PRT; 3130 AA.

O9BK46;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Reticulocyte binding protein 2 homolog A.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21101060; PubMed=11160005;
RX DOI=10.1128/TAI.69.2.1084-1092.2001;
RA Triglia T., Thompson J., Caruana S.R., Delorenzi M., Speed T.,
RA Cowman A.F.;
RT "Identification of proteins from Plasmodium falciparum that are
RT homologous to reticulocyte binding proteins in Plasmodium vivax.";
RL Infect. Immun. 69:1084-1092(2001).
DR EMBL: AF312916; AAK19244.1; -; Genomic DNA.
SQ SEQUENCE 3130 AA; 370419 MW; 13D973DB89D82026 CRC64;
Query Match 8.8%; Score 181.5; DB 2; Length 3130;
Best Local Similarity 19.2%; Pred. No. 0.22;
Matches 95; Conservative 111; Mismatches 137; Indels 153; Gaps 23;
QY 11 DKISSLELYKYIFPRNYIENVAEDCLKNGLILESAA-----HNSEVELARLVQV-LKNA 65
Db 535 EKASMDENEMKKELSLNNYIEK-TDYILQTYNIFKSKNIIINNNSKNISSKYITIEGLKND 593
QY 66 L-LNCIISYRFGHIGYVLTOTLIDLEQVNIELPIGFEYL-----YEVVRLDGLVD 118
Db 594 IDELSLSISY-----FKDSQETLIKDEKKNMKTDYLNKYNKYIENVTINEIILL 645
QY 119 PDHITYKVSNNKNSLDVAKHKSRLLIYENFDYILKRYVPCYTESFLDLYLPKRYV 178
Db 646 KDSITQRIADIDELNSLININ-----DFINEKNISQEKVSYNLN-KLYKGSFE 694
QY 179 EIERRIENH-----NPLFYKDESLVOLQDALSATTSLTOSNDRSGSILSFLRKONS 235
Db 695 ELESLSHFLPTKYLTFIEKKSQVNELOTILNNSNECAK-----NFMKSDNN 741
QY 236 NNHSK-----DISNRLNDSLSOELA-RUKSNLNEGMYFTATPS----- 275
Db 742 NNNNSNINILKTELSHLLSKENIKKLNHIEQIONSNTKITTYTIDNNEMDYK 801
QY 276 ---ASLEVIKYDLSYL-----KEALALIKAKIGADTK 304
Db 802 EEIESLEVYKHTIGNIQKEYILHLVENDKNALAVHTSMQILOKDAIQNKIKNSDDIK 861
QY 305 EPLTRSFNQAKGLNGKGRSNYDPL-KGVQE-----QVENACNLKTKYF----- 352
Db 862 ---ILKKYKEMNQ-----DLLNYYEILDKGLKONTYIKEMHTASLVQITQIPIYEDKT 911
QY 353 ---GLDMKFNLSLSEKQVE---RDIKLIIEYSK-----YNO 385
Db 912 ISELEQEPN-----NNNQKLDNILOQINAMNLNINLOTLGINACNTNNNNVHEHLNKK 966
QY 386 LIQSSSFNDEELAILK 401
Db 967 KIELKNILNDQMKIK 982
RESULT 15
Q9BK46_PLAF7
ID Q9BK46_PLAF7 PRELIMINARY; PRT; 3130 AA.
AC Q9BK46;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 10-MAY-2005 (TREMBLrel. 30, Last annotation update)
DE Reticulocyte binding protein 2 homolog a (Normocyte binding protein
DE 2a).
GN Name=PF13_0198;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:30:59 ; Search time 25.0907 Seconds
(without alignment)
1341.095 Million cell updates/sec

Title: US-10-688-058-4
Perfect score: 2062
Sequence: 1 MCDLRKTKLIDKISLELYK.....OSSSFDNEELAILKEKLFSP 407

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgm2_6/ptodata/1/iaa/5 COMB pep.*
2: /cgm2_6/ptodata/1/iaa/6 COMB pep.*
3: /cgm2_6/ptodata/1/iaa/H COMB pep.*
4: /cgm2_6/ptodata/1/iaa/PCTUS COMB pep.*
5: /cgm2_6/ptodata/1/iaa/RE COMB pep.*
6: /cgm2_6/ptodata/1/iaa/backfiles1 pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	157.5	7.6	583	US-09-710-279-1358	Sequence 1358, Ap
2	157.5	7.6	1010	US-09-134-001C-5178	Sequence 5178, Ap
3	153.5	7.4	448	US-09-328-352-6856	Sequence 6856, Ap
4	136	6.6	2662	US-09-595-684B-31	Sequence 31, Appl
5	133.5	6.5	2663	US-09-538-092-1252	Sequence 1252, Ap
6	132.5	6.4	1780	US-09-949-016-6899	Sequence 6899, Ap
7	132.5	6.4	1786	US-09-949-016-7880	Sequence 7880, Ap
8	132	6.4	511	US-09-248-796A-16380	Sequence 16380, A
9	131.5	6.4	878	US-09-134-001C-4378	Sequence 4378, Ap
10	130.5	6.3	1073	US-09-541-782-6	Sequence 6, Appl
11	130.5	6.3	1073	US-09-723-820-6	Sequence 6, Appl
12	130.5	6.3	1073	US-10-270-085-6	Sequence 6, Appl
13	129.5	6.3	575	US-09-724-519-8	Sequence 8, Appl
14	129.5	6.3	575	US-09-592-037-8	Sequence 8, Appl
15	129.5	6.3	575	US-09-428-156B-8	Sequence 8, Appl
16	129.5	6.3	905	US-09-248-796A-16333	Sequence 16333, A
17	129	6.3	1241	US-09-248-796A-18397	Sequence 18397, A
18	127.5	6.2	367	US-09-543-681A-8041	Sequence 8041, Ap
19	126	6.1	5024	US-09-710-279-2964	Sequence 2964, Ap
20	125.5	6.1	435	US-09-762-194-8	Sequence 8, Appl
21	125.5	6.1	435	US-09-762-194-10	Sequence 10, Appl
22	123.5	6.0	415	US-09-543-681A-8002	Sequence 8002, Ap
23	123	6.0	630	US-09-248-796A-20275	Sequence 20275, A
24	123	6.0	3878	US-09-914-259-11	Sequence 11, Appl
25	121	5.9	872	US-08-851-843A-8	Sequence 8, Appl
26	121	5.9	872	US-08-851-843A-54	Sequence 54, Appl
27	121	5.9	872	US-08-974-549A-221	Sequence 221, Appl

28	121	5.9	872	2	US-08-854-050-8	Sequence 8, Appl
29	121	5.9	872	2	US-08-854-050-54	Sequence 54, Appl
30	121	5.9	872	2	US-09-430-323-8	Sequence 8, Appl
31	121	5.9	872	2	US-09-430-323-54	Sequence 54, Appl
32	121	5.9	872	2	US-09-402-181B-221	Sequence 221, Appl
33	121	5.9	872	2	US-09-721-456-221	Sequence 221, Appl
34	121	5.9	872	2	US-09-766-253-8	Sequence 8, Appl
35	121	5.9	872	2	US-09-766-253-54	Sequence 54, Appl
36	121	5.9	872	2	US-10-054-235-8	Sequence 8, Appl
37	121	5.9	872	2	US-10-054-235-54	Sequence 54, Appl
38	121	5.9	872	2	US-09-438-486A-8	Sequence 8, Appl
39	121	5.9	872	2	US-09-438-486A-54	Sequence 54, Appl
40	121	5.9	10182	2	US-09-134-001C-3159	Sequence 3159, Ap
41	119.5	5.8	1078	2	US-09-248-796A-20284	Sequence 20284, A
42	119	5.8	905	2	US-09-248-796A-15706	Sequence 15706, A
43	118	5.7	448	2	US-09-248-796A-20279	Sequence 20279, A
44	118	5.7	722	2	US-09-248-796A-20613	Sequence 20613, A
45	118	5.7	1427	2	US-09-538-092-1044	Sequence 1044, Ap

ALIGNMENTS

RESULT 1
US-09-710-279-1358
; Sequence 1358, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1358
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-1358

Query Match	7.6%	Score 157.5;	DB 2;	Length 583;
Best Local Similarity	25.6%	Pred. No. 4.2e-06;		
Matches	80;	Conservative 55;	Mismatches 98;	Indels 79; Gaps 16;
QY	131	KNNSLDVAKIHKSRLLIYENFDYILKRVPCYTESFLDIYLFKEIYVEIERR----	IEN- 186	
Db	154	QNSKDKQSILRT-LFNSERFDEI-----RHLLVENVKQEK--VOIENRYQIENL 201		
QY	187	HNFLFYDESLVQLQDALSSATTSLSALTOSNDRGSGILSSFLRKQNSNNHSDIENL 245		
Db	202	WNIDITFNNDLALYKELESSQTDKMIKFPQFNDYCKILKSP--EBAKKITKELDDL 259		
QY	246	RN--LNDLSLQELARKSNLNNEGMFYTATPSASLEVIKYD-----LSYLKEALAL 294		
Db	260	NHKYKNVSELSENTKLLA-----EKIKFDLLKKEQNYDKLQKELKM 302		
QY	295	I---KAKIGADTKBPLTRSFNEQAKGLGNDGKDRSNFYDFLKGVOEQVENACN----- 345		
Db	303	IQESKVLITYTRLSQSKKDKDELVSLEHQSKLNETVYHNEIKGFKQKQLEHLSTRENEIT 362		
QY	346	-----LKUTKYPGLDMKFNSL--IMLSEOK--VERDILKLIELSKYKNQLI-----QS 389		
Db	363	QFNOYLEKNQVF-----FNQLDKIISYQKQKPVIEEIK--RLYSEYNDLITKKEELTKE 415		
QY	390	SSFDNEELAILK 401		
Db	416	MNNKKNKQFAIIE 427		

RESULT 2
 US-09-134-001C-5178
 ; Sequence 5178, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134,001C
 ; PRIORITY FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 5178
 ; LENGTH: 1010
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-5178

Query Match 7.6%; Score 157.5; DB 2; Length 1010;
 Best Local Similarity 25.6%; Pred. No. 9.9e-06;
 Matches 80; Conservative 55; Mismatches 98; Indels 79; Gaps 16;
 QY 131 KNSLDVAKHKSRLLIYENFYILKRYVPCYTESFLDIYLFEXIYVEIER--IEN- 186
 DB 162 QNSKDKQSILRT-LFNSRPFDEI-----RHLLVENVKQEK--VQIENRYTOIENL 209
 QY 187 -HNFLFYKDESIVQLODALSSATTSLSALTQSNDRGSGILSSFLRKQNSNNHSDINL 245
 DB 210 WNDITFNNDLALYKLESSTQDKMIEKFPQFNDYGCCKILKSP--BEAKNKITKELDDL 267
 QY 246 RN---LNDSLSQELARLKSNNLNEGMYTATPSASLEVIKPD-----LSYLKEALAL 294
 DB 268 NHKYKVNVELSENKYLKA-----EKIFDILKEQNYIDKLKQELKM 310
 QY 295 I---KAKIGADTKPLTRSFNEQAKGLGNDGKDRSNYYDFLKGVOEQVENACN----- 345
 DB 311 IQESKVLITYFRLQSLKDKDELVSLEHQSKLNETVYHNEIKFKQKLEHLSTRENEIT 370
 QY 346 -----IKLTKYGLDMKNSL--IMLSEOK--VERDIKLIELYKYNQLI-----QS 389
 DB 371 QFNQYLEKNQVP-----FNQLDKIISYQKQPVIEEEK--RLYSEYNDLITKKEELTKB 423
 QY 390 SSFNDNEELAILK 401
 DB 424 MNKKNDFAIE 435

RESULT 3
 US-09-328-352-6856
 ; Sequence 6856, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; PRIORITY FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 6856
 ; LENGTH: 448
 ; TYPE: PRT
 ; ORGANISM: Acinetobacter baumannii
 US-09-328-352-6856
 Query Match 7.4%; Score 153.5; DB 2; Length 448;
 Best Local Similarity 21.8%; Pred. No. 6.6e-06;

Matches 87; Conservative 59; Mismatches 145; Indels 109; Gaps 18;
 QY 26 RNYIENVAEDCLKNGLILLESAA-HNVSEVELARLKVQKNALLNCIIISVRPHGICGVLVK 84
 DB 59 KRVNRPASDMLRAGWFYEGIQDRDLIKLEACAKAFQNLQILLSSLLISRLYGVVYVLIG 118
 QY 85 TKDTLIDLEQPVNI-ELPIG----FEYLDYVYVRDLGVDFDHIT-----123
 DB 119 TVDG-GNLEQFPDLKGLGVRLEFFVLKKYIKP-----DTTIVLSPKVCGLLKQPEF 172
 QY 124 YKVSNNKNSLDVAKHKSRLLIYENFYILKRYVPCYTESFLDIYLFEXIYVEIER 183
 DB 173 YKLQMDGKAPQ---RIHHSRLIKFCHADVNEE-----POSILQEVY-----211
 QY 184 IENHNFLFYKDESIVQLODALSSATTSLSALTQSNDRGSGILSSFLRKQNSNNHSHK-DI 242
 DB 212 -----EDLLDHAHV-----KGSASLV-----HESKIDV 235
 QY 243 SNLRNLNDSLSQELARLKSNNLNEGMYTATPSASLEVIK-----YDLSYLKEALALIK 296
 DB 236 IKTPGLVDKIKEDMKVAERPLSVGLLKGNGMIVLDAEEDYDSKTYNFAGLPDMRRFS 295
 QY 297 AKIGADTKPLTRSFNEQAKGLGNDGKDRSNYYD-----FLKGVOEQV-ENACNLK 347
 DB 296 IOTAGAADIPYTLIFGQSPAGMNATGEHDTNRNYDSIATKQTMWLKPFMMQILDVICOTT 355
 QY 348 LTKYP-GLDMKFNSLIML-----SEEQV--ERDIKLIEL 379
 DB 356 FGRVFPNLDIVFNPLWQLDAKVRSEVEKANSEKARDAKYLEM 395

RESULT 4
 US-09-595-684B-31
 ; Sequence 31, Application US/09595684B
 ; Patent No. 6544766
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Ohashi, Cara
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Vaisberg, Eugeni
 ; APPLICANT: Wood, Kenneth
 ; APPLICANT: Yu, Ming
 ; TITLE OF INVENTION: Human kinesins and methods of producing
 ; FILE REFERENCE: cytop036
 ; CURRENT APPLICATION NUMBER: US/09/595,684B
 ; PRIORITY FILING DATE: 2002-06-24
 ; PRIOR APPLICATION NUMBER: 09/295,612
 ; PRIOR FILING DATE: 2000-04-20
 ; NUMBER OF SEQ ID NOS: 105
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 31
 ; LENGTH: 2662
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-595-684B-31
 Query Match 6.6%; Score 136; DB 2; Length 2662;
 Best Local Similarity 22.1%; Pred. No. 0.005;
 Matches 97; Conservative 80; Mismatches 153; Indels 108; Gaps 21;
 QY 1 MCDLRKTKLIDKISSL--ELYKYSIFFRNYIENVAEDCLKNGLILLESAAHNV-SEVELAR 57
 DB 713 LCNL---ELEGIITDLQKELNK-----EVEENALREVEILLSELKLPSEVERUR 760
 QY 58 LKVLQKNALLNCIIS-----YRFHGIGYVLVTKDTLIDLEQPVNIELPIGF 104
 DB 761 KEIQDKSEELHIITSEKDKLFEVVKESRVQGLEEGTKKDDLATQOS-----NY 812
 QY 105 EYLDYEV--VRDLGVDFPHITYKVKNSNNKNSLDAVKIHKRLIYENPDY---ILKRYV 159
 DB 813 KSTDQEFQNFKTLHMDFEQYKVMLEENRMNQEIIVNLSKEA-----QKFDSSLGALKTEL 868


```
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7880
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7880

Query Match      6.4%; Score 132.5; DB 2; Length 1786;
Best Local Similarity 21.7%; Pred. No. 0.0057;
Matches 90; Conservative 63; Mismatches 143; Indels 119; Gaps 17;

QY 11 DKISLELYKYSI-----PFRNVIENVAEDCLNGLLILESAHNVSEVE--LARLKVQ 61
DB 998 DYLVSKQVKEVRIQPNRNSPHSSIEAIWEECKEIVKASSKSHQIELEQOIEKQAE 1057
QY 62 LKNAL-LNCIIISYRPHGIGYVLVKTDTLI-----DLEQPVNIEIPGFEYLDYEVVDLG 116
DB 1058 VAGYKDENNRLKEKEHKQDDLLKEKETLIQOLKEELQKVTLDVQIQHV-VEGKRLS 1116
QY 117 VDFDHT-YKVKNNKNSLDAVKIHKRLIYENPDYILKRYVPCYTESFLDIDYLPFK 175
DB 1117 ELTQGVTCYKAKIKETITLETQKVERSHSAKLEQ-----DILEKES 1158
QY 176 IVEIETRIENHNFLFYKDESIVQLODALSATTSLSALTQSNDRGSGILSSFLRQK-- 233
DB 1159 IILKLERNLK-----EFOEHQDSV-----KNTKDLNVKELKKEBIT 1196
QY 234 NNNHSHKDISNRLN-----NDLSLQELARLKNLNEGMYTATPSASLEVIKYDLSYLK 289
DB 1197 QLTNQLQDKHLLQLKESEETNRQETEKKEEL-----SASSARTQNLKADLQKRE 1248
QY 290 EALALIKAKIGADTKPELTRSFNEQAKGLNGDKGRSNYYDFLKGVEQVENACNLKLT 349
DB 1249 EDY-----ADLKEKLT-----DAKKQ-----IKQVQKEVS----- 1273
QY 350 KYFGLDMKFNLSLMLSEQKVERDIKLIYLSKYNQLOIQQSSFDNEELAILKEKL 404
DB 1274 -----VMRDEKLLR-IXINELEKKNQCSQELDMKQRTIQQLKBQL 1314

RESULT 8
US-09-248-796A-16380
; Sequence 16380, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16380
; LENGTH: 511

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7880
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7880

Query Match      6.4%; Score 132.5; DB 2; Length 1786;
Best Local Similarity 21.7%; Pred. No. 0.0057;
Matches 90; Conservative 63; Mismatches 143; Indels 119; Gaps 17;

QY 11 DKISLELYKYSI-----PFRNVIENVAEDCLNGLLILESAHNVSEVE--LARLKVQ 61
DB 998 DYLVSKQVKEVRIQPNRNSPHSSIEAIWEECKEIVKASSKSHQIELEQOIEKQAE 1057
QY 62 LKNAL-LNCIIISYRPHGIGYVLVKTDTLI-----DLEQPVNIEIPGFEYLDYEVVDLG 116
DB 1058 VAGYKDENNRLKEKEHKQDDLLKEKETLIQOLKEELQKVTLDVQIQHV-VEGKRLS 1116
QY 117 VDFDHT-YKVKNNKNSLDAVKIHKRLIYENPDYILKRYVPCYTESFLDIDYLPFK 175
DB 1117 ELTQGVTCYKAKIKETITLETQKVERSHSAKLEQ-----DILEKES 1158
QY 176 IVEIETRIENHNFLFYKDESIVQLODALSATTSLSALTQSNDRGSGILSSFLRQK-- 233
DB 1159 IILKLERNLK-----EFOEHQDSV-----KNTKDLNVKELKKEBIT 1196
QY 234 NNNHSHKDISNRLN-----NDLSLQELARLKNLNEGMYTATPSASLEVIKYDLSYLK 289
DB 1197 QLTNQLQDKHLLQLKESEETNRQETEKKEEL-----SASSARTQNLKADLQKRE 1248
QY 290 EALALIKAKIGADTKPELTRSFNEQAKGLNGDKGRSNYYDFLKGVEQVENACNLKLT 349
DB 1249 EDY-----ADLKEKLT-----DAKKQ-----IKQVQKEVS----- 1273
QY 350 KYFGLDMKFNLSLMLSEQKVERDIKLIYLSKYNQLOIQQSSFDNEELAILKEKL 404
DB 1274 -----VMRDEKLLR-IXINELEKKNQCSQELDMKQRTIQQLKBQL 1314

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7880
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7880

Query Match      6.4%; Score 132.5; DB 2; Length 1786;
Best Local Similarity 21.7%; Pred. No. 0.0057;
Matches 90; Conservative 63; Mismatches 143; Indels 119; Gaps 17;

QY 11 DKISLELYKYSI-----PFRNVIENVAEDCLNGLLILESAHNVSEVE--LARLKVQ 61
DB 998 DYLVSKQVKEVRIQPNRNSPHSSIEAIWEECKEIVKASSKSHQIELEQOIEKQAE 1057
QY 62 LKNAL-LNCIIISYRPHGIGYVLVKTDTLI-----DLEQPVNIEIPGFEYLDYEVVDLG 116
DB 1058 VAGYKDENNRLKEKEHKQDDLLKEKETLIQOLKEELQKVTLDVQIQHV-VEGKRLS 1116
QY 117 VDFDHT-YKVKNNKNSLDAVKIHKRLIYENPDYILKRYVPCYTESFLDIDYLPFK 175
DB 1117 ELTQGVTCYKAKIKETITLETQKVERSHSAKLEQ-----DILEKES 1158
QY 176 IVEIETRIENHNFLFYKDESIVQLODALSATTSLSALTQSNDRGSGILSSFLRQK-- 233
DB 1159 IILKLERNLK-----EFOEHQDSV-----KNTKDLNVKELKKEBIT 1196
QY 234 NNNHSHKDISNRLN-----NDLSLQELARLKNLNEGMYTATPSASLEVIKYDLSYLK 289
DB 1197 QLTNQLQDKHLLQLKESEETNRQETEKKEEL-----SASSARTQNLKADLQKRE 1248
QY 290 EALALIKAKIGADTKPELTRSFNEQAKGLNGDKGRSNYYDFLKGVEQVENACNLKLT 349
DB 1249 EDY-----ADLKEKLT-----DAKKQ-----IKQVQKEVS----- 1273
QY 350 KYFGLDMKFNLSLMLSEQKVERDIKLIYLSKYNQLOIQQSSFDNEELAILKEKL 404
DB 1274 -----VMRDEKLLR-IXINELEKKNQCSQELDMKQRTIQQLKBQL 1314
```

```
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16380

Query Match      6.4%; Score 132; DB 2; Length 511;
Best Local Similarity 20.2%; Pred. No. 0.0009;
Matches 92; Conservative 72; Mismatches 149; Indels 142; Gaps 22;

QY 49 NVSEVELARLVQVKNALLNCIIISYRPH-GIGYVLVK-TKDTLIDLEQPVNIELPIGFEY 106
DB 41 DIHEKRMQR-NIEIKSILQITMYAKLNERVDYLLEKLTSTELDSEKVM-----K 91
QY 107 LDYEVYRDLGVDFHITYKVKNNKNSLDAVKIHKSLIYENPDYILKRYVPCYTESF 166
DB 92 LNSEFDPQKEFDYDKL-IKDKGLTLRKGLKDLKFDREI---EN-----TPCYNE-- 137
QY 167 LLDIYLFKEIYVEI-----ERRIENHNFLFYKDSLVQLQDALSATTSLSALTOS---- 217
DB 138 ----MIEDLFVQIKDDHPETKTGDGLIEYLKEHRNRIDDLVLSKQTIKLDLLYQKAO 192
QY 218 --NDRGSGILSSFLRKQNSNNHSHKDIS-----NL-----RNL 248
DB 193 IVSDDLHTGDFDRSFLNKDKPEEKENLDKPKAPAKTKTVTVTETINLPKPVENDTKEI 252
QY 249 NDSL-----SOELARLKS-NLNNEGMYTATPSASLEVIK-----YD----- 284
DB 253 LDELEILPATKEPAKIPSDNLSKAAEFILKHPISICTEQOKDALIMTAFDQLENKSDK 312
QY 285 -----LSYLKE-----ALALIKAKIGADTKPELTRSFNFOAKGLGNDG 322
DB 313 HIVHQSLLLVQVQLSGNGKAPANVINAIKLFPSKIAAES-SPAKHAFLE----- 362
QY 323 KGDRSNYYDFLKG-----VQEOVENACNLK-----LTKYFGLDMKFNLSLMLSE 368
DB 363 --DVNQTFNHIKGRCEIKEEQRNAANKGTGEBEEBALIQLKALDDNTTELLVNIPOEG 420
QY 369 KVERDIKLIYLSKYNQLOIQQSSFD--NEELAILK 401
DB 421 TKEVEIFTKLPIEFQNAIKTESIDENVKEPAKLK 455

RESULT 9
US-09-134-001C-4378
; Sequence 4378, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4378
; LENGTH: 878
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4378

Query Match      6.4%; Score 131.5; DB 2; Length 878;
Best Local Similarity 23.4%; Pred. No. 0.0023;
Matches 94; Conservative 79; Mismatches 144; Indels 85; Gaps 25;

QY 27 NVIEN-VAEDCLNGLLILESAHNVSEVELARLVQVKNALLNCI----- 70
DB 363 NNVSNYTVKDLPN---IEVYSQN-----EIIDLTN-ENAKLINLRFLDKDRNDKKE 414
QY 71 -ISYRPHGIGYVLVKTDTLIDLEQPVNIELPIGFEYLDYEVYRDLGVDFDHTYKVKNS 129
```

Db 415 EIKTNLHNSKSLIKAKEDLENLQSKIN-QLPKLEKL--KHFNELG-----IGKKLVQ 466

QY 130 NKNNSLDAVKIHKSLRIIYENFDYILKRYVPCYTESFLDDIYLFPEKIYVEIERRIENHNF 189

Db 467 GK-ISREEQYQNTQKIIEDN-DISITNIIIPFNENYQIKHVE-IFDSIKNIIDNEN- 522

QY 190 LFYKDESIVQODALSSATTSLSALTQSNDRGSGILSSFLRKQNSNNHSDISN-LRNL 248

Db 523 -----KLUKE-IKSMFTDLKOTTONIEIKIYNVW-----KEKKKNIKEINRAIKSL 568

QY 249 NDSLSQELARUKSNLNNMGFTATPS--ASLEVIKYDLISYLKEALALIKAKIGADTKEP 306

Db 569 DDI-----EGTKEDIAHE---YETOKQITSEPLETQLSRVKTSITIELENE-RIQLKED 620

QY 307 LTRSFNEQAKGLGNDGKDRSNYYDFLKGVOQVENACNLKLTIFYGLDMKFNLSLIMSE 366

Db 621 LKEIFEQLKNLRCVKINRY-----LKKQV-----NIQIPYANVN---NLIEFLKE 667

QY 367 EOKV-ERDIKLIYLSKYN-----QLIOSSFDNEELAILKE 402

Db 668 ENGLGDSTLKWIKNHQSFNFPKFIKLIK-----DRDSEAIYEE 705

RESULT 10

US-09-541-782-6

Sequence 6, Application US/09541782

Patent No. 6284480

GENERAL INFORMATION:

APPLICANT: Nislow, Corey

APPLICANT: Sakowicz, Roman

APPLICANT: Beraud, Christophe

TITLE OF INVENTION: Antifungal Assay

FILE REFERENCE: 1015

CURRENT APPLICATION NUMBER: US/09/541,782

CURRENT FILING DATE: 2000-04-03

NUMBER OF SEQ ID NOS: 10

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 6

LENGTH: 1073

TYPE: PRT

ORGANISM: S. pombe

US-09-541-782-6

Query Match 6.3%; Score 130.5; DB 2; Length 1073;

Best Local Similarity 19.5%; Pred. No. 0.004;

Matches 85; Conservative 68; Mismatches 167; Indels 115; Gaps 15;

QY 30 ENVAEDCLKNGLILESAAHNVSEVELARLKVKLKNALLNCI--ISYRFHGIGYVLV--- 83

Db 320 ENIGRGAENKRAETGMINOSLLTGR-----VINALVEKAHHIYRESKLTRLLODSLG 375

QY 84 -KTQDITLIDLEQPVNIELPIGPEYLDYE-----YVRDLGVDFDHI 122

Db 376 GKTKTSMIVTVSSVTNLEETISTLEYAARAKSIRKPNQNLVFRKVLIKDLVLDIERL 435

QY 123 TVYKSNNNKNSLDVAKIHKSLRIIYENFDYILKRYVPCYTESFLDDIYLFPEKIYVEIER 182

Db 436 KNDLNATRKNG-----VYLAESTYKELMD 460

QY 183 RIENHNFYKDESIVQLODA-LSSATTSLSALTQSNDRGSGILSSFLRKQNSNNHSD 241

Db 461 RVQNKDLLCQEQARKLEVLDLNVKSSREQLQVYVSKNQHKEVEALQQLVNS---STE 517

QY 242 ISNLRNLNDSLSQELARUKSNLNNMGFTATPSASLEVIKYDLSYL---KEALALIKA 297

Db 518 LESVKSENEKLNELV-----LEIEKRKKYETNEAKITTVATDLSQYVRESKEYIASIYE 572

QY 298 KIG---ADTKE-----PLTSEFNEQAKGLGNDGKDRSNYY-----DPLKGV 336

Db 573 KLDRTERRNNKENNFNWLKFNLLTMLRSFHSFT-----DETNGYFTLLDNFNASM 624

QY 337 QBOVENACN-----LKLTKYF-GLDMKFNLSLIMLSEEQKVERDIKLIYLSKYNQLIQS 389

Db 625 BELNTHSNQLLISMTKITEHFQSLDEALQASRSCAVPNSSLDLIIVSELKDKSNLSLDA 684

QY 390 SSFNEELAILKEKL 404

Db 685 LEHSLQDISMSQKL 699

RESULT 12

US-10-270-085-6

Sequence 6, Application US/10270085

Patent No. 6627408

GENERAL INFORMATION:

APPLICANT: Nislow, Corey

APPLICANT: Sakowicz, Roman

Db 625 BELNTHSNQLLISMTKITEHFQSLDEALQASRSCAVPNSSLDLIIVSELKDKSNLSLDA 684

QY 390 SSFNEELAILKEKL 404

Db 685 LEHSLQDISMSQKL 699

RESULT 11

US-09-723-820-6

Sequence 6, Application US/09723820

Patent No. 6468760

GENERAL INFORMATION:

APPLICANT: Nislow, Corey

APPLICANT: Sakowicz, Roman

APPLICANT: Beraud, Christophe

TITLE OF INVENTION: Antifungal Assay

FILE REFERENCE: 1015

CURRENT APPLICATION NUMBER: US/09/723,820

CURRENT FILING DATE: 2000-11-28

PRIOR APPLICATION NUMBER: 09/541,782

PRIOR FILING DATE: 2000-04-03

NUMBER OF SEQ ID NOS: 10

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 6

LENGTH: 1073

TYPE: PRT

ORGANISM: S. pombe

US-09-723-820-6

Query Match 6.3%; Score 130.5; DB 2; Length 1073;

Best Local Similarity 19.5%; Pred. No. 0.004;

Matches 85; Conservative 68; Mismatches 167; Indels 115; Gaps 15;

QY 30 ENVAEDCLKNGLILESAAHNVSEVELARLKVKLKNALLNCI--ISYRFHGIGYVLV--- 83

Db 320 ENIGRGAENKRAETGMINOSLLTGR-----VINALVEKAHHIYRESKLTRLLODSLG 375

QY 84 -KTQDITLIDLEQPVNIELPIGPEYLDYE-----YVRDLGVDFDHI 122

Db 376 GKTKTSMIVTVSSVTNLEETISTLEYAARAKSIRKPNQNLVFRKVLIKDLVLDIERL 435

QY 123 TVYKSNNNKNSLDVAKIHKSLRIIYENFDYILKRYVPCYTESFLDDIYLFPEKIYVEIER 182

Db 436 KNDLNATRKNG-----VYLAESTYKELMD 460

QY 183 RIENHNFYKDESIVQLODA-LSSATTSLSALTQSNDRGSGILSSFLRKQNSNNHSD 241

Db 461 RVQNKDLLCQEQARKLEVLDLNVKSSREQLQVYVSKNQHKEVEALQQLVNS---STE 517

QY 242 ISNLRNLNDSLSQELARUKSNLNNMGFTATPSASLEVIKYDLSYL---KEALALIKA 297

Db 518 LESVKSENEKLNELV-----LEIEKRKKYETNEAKITTVATDLSQYVRESKEYIASIYE 572

QY 298 KIG---ADTKE-----PLTSEFNEQAKGLGNDGKDRSNYY-----DPLKGV 336

Db 573 KLDRTERRNNKENNFNWLKFNLLTMLRSFHSFT-----DETNGYFTLLDNFNASM 624

QY 337 QBOVENACN-----LKLTKYF-GLDMKFNLSLIMLSEEQKVERDIKLIYLSKYNQLIQS 389

Db 625 BELNTHSNQLLISMTKITEHFQSLDEALQASRSCAVPNSSLDLIIVSELKDKSNLSLDA 684

QY 390 SSFNEELAILKEKL 404

Db 685 LEHSLQDISMSQKL 699

RESULT 12

US-10-270-085-6

Sequence 6, Application US/10270085

Patent No. 6627408

GENERAL INFORMATION:

APPLICANT: Nislow, Corey

APPLICANT: Sakowicz, Roman

; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/10/270,085
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/723,820
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/541,782
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1073
; TYPE: PRT
; ORGANISM: S.pombe
US-10-270-085-6

Query Match 6.3%; Score 130.5; DB 2; Length 1073;
Best Local Similarity 19.5%; Pred. No. 0.004;
Matches 85; Conservative 68; Mismatches 167; Indels 115; Gaps 15;

QY 30 ENVAEDCLKNGILLESAAHNVSEVELARLKVOLKNALLNCI--ISVRPHGIGYVLV---- 83
Db 320 ENIRSGAENKARETGMINQSLLTIGR---VINALVEKAHHIPYRESKLTRELQDSLG 375
QY 84 -KTKDTLIDLEQPVNIELPIGPEYLDYE-----YVRDLGVDFDHI 122
Db 376 GKTKTSMIVTGSTNTNLEETISTLEYAARAKSIRKPNQNLVFRKVLKDLVDLIERL 435
QY 123 TVYKSNNNKNSLDAVKIHKSLIIYENFDYILKRVPCYTESFLDYLFPFKIYVEIER 182
Db 436 KNDLNATRKNG-----VYLAESTYKELMD 460
QY 183 RIENHNFYKDESLOLQDA-LSSATTSLSALTQSNDRGSGILSSFLRKQNSNNHSD 241
Db 461 RVQNDLLCQEQARKLEVLDLVKVSREQLQYVSKNSQHEKVEALQQLVNS---STE 517
QY 242 ISNRLNLSLSOELARLKSNNNEGMFTATPSASLEVIKYDLSYL---KEALALIKA 297
Db 518 LESVKSENEKLNELV-----LEIEKRYKYETNEAKITVATDLSQYRESKEYIASLYE 572
QY 298 KIG---ADTKE-----PLTRSFNQAKGLGNDGKGRSNYY-----DFLKGV 336
Db 573 KLDRTERNKENENNFWNLKFNLLTMLRSFHGSFT-----DETNGYFTLLDNFNASM 624
QY 337 QOVENACH-----LKLTKYP-GLDMKFNLSLIMSEQKVERDIKLIELYSKYNLIQS 389
Db 625 EELLNTHSNQLISMKTITEHPQSLDEALQSARSSCAVPNSSLDLIVSELKDSKNSLLDA 684
QY 390 SSFDEEELAILKEKL 404
Db 685 LEHSIQDISMSQKL 699

RESULT 13

US-09-724-519-8
; Sequence 8, Application US/09724519
; Patent No. 6414121
; GENERAL INFORMATION:
; APPLICANT: Wood, Kenneth
; APPLICANT: Finer, Jeffrey
; APPLICANT: Beraud, Christophe
; APPLICANT: Mak, John
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: Methods of screening for modulators of
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell
; TITLE OF INVENTION: proliferation states
; FILE REFERENCE: 1014A
; CURRENT APPLICATION NUMBER: US/09/724,519
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/592,037
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 09/428,156

; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Human
US-09-724-519-8

Query Match 6.3%; Score 129.5; DB 2; Length 575;
Best Local Similarity 22.4%; Pred. No. 0.0019;
Matches 105; Conservative 63; Mismatches 175; Indels 125; Gaps 22;

QY 7 TKLIDKISSLELYKYSIF-FRNYIENVAEDCL-----KNGILIL-----ESAAHNVSEV- 53
Db 152 TEFSVKVSLLEIYNEELFDLLNPSDSVSERLQMPDPRKRGVILKGLLEIIVHNKDEVY 211
QY 54 ---ELARLKVOLKNALLNCIISYRPHGIGYVLVTKOTLIDLEQPVNIELPIGFYLDYE 110
Db 212 GILEKGAARCTTAATLMNA-YSSRSHSVFVTIHKMETTIDGEEELVKIG---KLNLDVLA 267
QY 111 YVRDLG----VDFDHIITYKVSNNKNSLDAVKIHKSLIIYENFDYILKR--YVPCYTE 164
Db 268 GSENIGRSGAVD-----KRAREAGNIQSLLTIGRVITALVERTPHVP-VRE 313
QY 165 SFLLDIYLFEXIYVEIERRIENHNFLFYKDESLOLQDALSAT-TLSALTQSNDRGS 223
Db 314 SKLIRI-----LQDSLGRFTSTIATISP-----AS 340
QY 224 GILSFLRKQNSNNHSDISNRLNLSL-----SOLARLKSML-----NNEGMPY 270
Db 341 LNLEETLSTLEVAHRAKNILNKPEVNQKLTKKALIKETIEERLKRDLAAAREKNGVYI 400
QY 271 TATPSASLEVIKYDLSYLKEALALIKAKIGA-----DTKEPL----- 307
Db 401 S---EENFRVMSGKLTVOEQIIVELIEKIGAVEEELNRVTELFMDNKNELDQCKSDLQNK 457
QY 308 TRSFNEQAKGLGNDGKGRSNYY--DFLKGVQEQVENACNLKL-----TK-YFLGDMKF 358
Db 458 TQELTETTKHGETKGLGLVKEEYTSALSTEEKLUHDAASKLLNVTETTKDVSGLUHKL 517
QY 359 NSLIMLSEQKVERDI---KLIELYSKYNLIQSLSFDEEELAILKEK 403
Db 518 DRKRAVDQHNAAEQDIFGKNLSLNFNNNEELIKDGSLEQTEGKLISEE 565

RESULT 14

US-09-592-037-8
; Sequence 8, Application US/09592037
; Patent No. 6437115
; GENERAL INFORMATION:
; APPLICANT: Wood, Kenneth
; APPLICANT: Finer, Jeffrey
; APPLICANT: Beraud, Christophe
; APPLICANT: Mak, John
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: Methods of screening for modulators of
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell
; TITLE OF INVENTION: proliferation states
; FILE REFERENCE: 1014A
; CURRENT APPLICATION NUMBER: US/09/592,037
; CURRENT FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 09/428,156
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Human
US-09-592-037-8

Query Match 6.3%; Score 129.5; DB 2; Length 575;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:21:03 ; Search time 20.3361 Seconds
(without alignments)
1966.133 Million cell updates/sec

Title: US-10-688-058-6

Perfect score: 470

Sequence: 1 LGFVPSYVPISSISLKG.....SIKNEWILVPLWKPWAFS 91

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_21.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*
- 9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	470	100.0	91	8	ADU98747
2	468	99.6	275	2	AAV19955
3	468	99.6	290	2	AAV19954
4	468	99.6	290	8	ADU98749
5	193	41.1	384	7	ABO72983
6	186.5	39.7	573	5	ABBS4794
7	179	38.1	309	9	ABM96406
8	161.5	34.4	577	7	ADC95590
9	158.5	33.7	569	4	AAU34910
10	158.5	33.7	571	7	ADH87483
11	157	33.4	575	5	ABF27731
12	153	32.6	273	6	ADBO7854
13	153	32.6	334	6	ADBO7856
14	153	32.6	352	6	ADBO7858
15	152.5	32.4	320	6	ADBO7852
16	152.5	32.4	575	5	ABF27730
17	152.5	32.4	575	5	ABF29826
18	152.5	32.4	575	8	ADV88795
19	152.5	32.4	575	8	ADV80048
20	152.5	32.4	575	8	ADV82185
21	147	31.3	349	4	AAU46688
22	147	31.3	349	6	ABM43207
23	138	29.4	312	4	AAU33717
24	138	29.4	312	6	ABU15624

25	138	29.4	318	7	ABO68427	ABO68427 Pseudomon
26	136	28.9	417	4	AAU40421	AAU40421 Propionib
27	136	28.9	417	6	ABM36940	ABM36940 Propionib
28	135.5	28.8	296	8	ADG32088	ADG32088 Mutant B
29	130	27.7	302	6	ADB11550	ADB11550 Alloiooc
30	127	27.0	356	7	ABO68411	ABO68411 Pseudomon
31	127	27.0	380	6	ABU21819	ABU21819 Protein e
32	126	26.8	427	7	ABO73023	ABO73023 Pseudomon
33	124.5	26.5	300	5	ABB48738	ABB48738 Listeria
34	124	26.4	303	6	ADB07978	ADB07978 Alloiooc
35	124	26.4	326	6	ADB07980	ADB07980 Alloiooc
36	108	23.0	314	6	ABU39563	ABU39563 Protein e
37	88	18.7	317	7	ABO82583	ABO82583 Pseudomon
38	87	18.5	333	6	ABM69068	ABM69068 Photorhab
39	86	18.3	338	7	ABO63655	ABO63655 Klebsiell
40	76.5	16.3	903	6	ABU49526	ABU49526 Protein e
41	76	16.2	361	7	ABO84191	ABO84191 Pseudomon
42	74	15.7	319	5	ABP60851	ABP60851 Buchnera
43	74	15.7	573	5	ABP39811	ABP39811 Staphyloc
44	74	15.7	573	8	ADS05795	ADS05795 Staphyloc
45	73.5	15.6	687	5	ABB92975	ABB92975 Herbicida

ALIGNMENTS

RESULT 1
ADU98747
ID ADU98747 standard; protein; 91 AA.
XX
AC ADU98747;
XX
DT 24-FEB-2005 (first entry)
XX
DE Borrelia burgdorferi antigenic polypeptide seqid 6.
XX
KW antibacterial; vaccine; immune stimulation; immunity; antigen;
KW DNA library.
XX
OS Borrelia burgdorferi.
XX
FN WO2004103269-A2.
XX
PD 02-DEC-2004.
XX
PF 17-OCT-2003; 2003WO-US033056.
XX
PR 18-OCT-2002; 2002US-0419401P.
XX
PA (MACR-) MACROGENICS INC.
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Sykes KF, Hale KS, Johnston SA;
XX
DR WPI; 2004-834155/82.
XX
N-PSDB; ADU98746.
PT Immunizing a subject against Borrelia burgdorferi infection comprises
PT providing to the subject at least one Borrelia antigen or its fragment.
PS Claim 27; SEQ ID NO 6; 121pp; English.
CC The invention describes a method of immunizing a subject comprising
CC providing to the subject at least one Borrelia antigen or its fragment to
CC induce an immune response. Also described are: an isolated polynucleotide
CC comprising a sequence having at least 17 contiguous nucleotides in common
CC with a sequence not given in the specification; an isolated polypeptide
CC having at least 5 consecutive amino acids of the sequence not given in
CC the specification; a vaccine composition comprising at least one Borrelia
CC antigen or at least one polynucleotide encoding a Borrelia antigen;
CC screening for at least one test polypeptide or test polynucleotide
CC encoding a polypeptide for an ability to produce an immune response;
CC preparing a vaccine; vaccinating a subject; treating a subject infected

CC with a pathogen; passive immunization; vaccination; detecting Herpesvirus
 CC ; and detecting antibodies to a herpesvirus. The method is useful in
 CC immunizing a subject against Borrelia burgdorferi infection. This is the
 CC amino acids sequence of a Borrelia burgdorferi polypeptide that may be
 CC useful in immunizing a subject against infection by a member of the
 CC Borrelia species.
 XX
 SQ Sequence 91 AA;

Query Match 100.0%; Score 470; DB 8; Length 91;
 Best Local Similarity 100.0%; Pred. No. 3.5e-50;
 Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQGFVVPSPVPISSISELKGKDKFKNMIGIDAGAGTQIVTEQALNYGLSKYELVPS 60

Db 1 LQGFVVPSPVPISSISELKGKDKFKNMIGIDAGAGTQIVTEQALNYGLSKYELVPS 60

QY 61 SESVMLASLDSSIKRNEWILVPLMKPHWAFS 91

Db 61 SESVMLASLDSSIKRNEWILVPLMKPHWAFS 91

RESULT 2

AAI19955
 ID AAY19955 standard; protein; 275 AA.

XX
 AC AAY19955;

XX 19-JUL-1999 (first entry)

XX B. burgdorferi antigenic protein, t810.aa.

XX Antigenic protein; vaccine; Lyme disease; infection; detection.

XX Borrelia burgdorferi.

XX WO9859071-A1.

XX 30-DEC-1998.

XX 18-JUN-1998; 98WO-US012718.

XX 20-JUN-1997; 97US-0050359P.

XX 22-JUL-1997; 97US-0053344P.

XX 22-JUL-1997; 97US-0053377P.

XX 03-SEP-1997; 97US-0057483P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (MEDI-) MEDIMMUNE INC.

XX Choi GH, Erwin AL, Hanson MS, Lathigra R;

XX WPI; 1999-189980/16.

XX N-PSDB; AAX61652.

XX New isolated Borrelia burgdorferi nucleic acids - used to develop

XX products for the diagnosis, prevention and treatment of diseases caused

XX by Borrelia, particularly Lyme disease.

XX Claim 12; Page 134; 275pp; English.

XX This sequence represents a Borrelia burgdorferi (Bb) protein of the

XX invention, which is suitable for use in a vaccine. The Bb polypeptides

XX can be used in vaccines for eliciting protective antibodies to members of

XX the Borrelia genus, particularly for the use against Lyme disease in

XX humans and animals. They can be used for preventing or attenuating an

XX infection caused by a member of the Borrelia genus. The products can also

XX be used for detection of members of the Borrelia genus

Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQGFVVPSPVPISSISELKGKDKFKNMIGIDAGAGTQIVTEQALNYGLSKYELVPS 60

Db 94 IQGFVVPSPVPISSISELKGKDKFKNMIGIDAGAGTQIVTEQALNYGLSKYELVPS 153

QY 61 SESVMLASLDSSIKRNEWILVPLMKPHWAFS 91

Db 154 SESVMLASLDSSIKRNEWILVPLMKPHWAFS 184

RESULT 3

AAI19954

ID AAY19954 standard; protein; 290 AA.

XX
 AC AAY19954;

XX 19-JUL-1999 (first entry)

XX B. burgdorferi antigenic protein, f810.aa.

XX Antigenic protein; vaccine; Lyme disease; infection; detection.

XX Borrelia burgdorferi.

XX WO9859071-A1.

XX 30-DEC-1998.

XX 18-JUN-1998; 98WO-US012718.

XX 20-JUN-1997; 97US-0050359P.

XX 22-JUL-1997; 97US-0053344P.

XX 22-JUL-1997; 97US-0053377P.

XX 03-SEP-1997; 97US-0057483P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (MEDI-) MEDIMMUNE INC.

XX Choi GH, Erwin AL, Hanson MS, Lathigra R;

XX WPI; 1999-189980/16.

XX N-PSDB; AAX61651.

XX New isolated Borrelia burgdorferi nucleic acids - used to develop

XX products for the diagnosis, prevention and treatment of diseases caused

XX by Borrelia, particularly Lyme disease.

XX Claim 12; Page 134; 275pp; English.

XX This sequence represents a Borrelia burgdorferi (Bb) protein of the

XX invention, which is suitable for use in a vaccine. The Bb polypeptides

XX can be used in vaccines for eliciting protective antibodies to members of

XX the Borrelia genus, particularly for the use against Lyme disease in

XX humans and animals. They can be used for preventing or attenuating an

XX infection caused by a member of the Borrelia genus. The products can also

XX be used for detection of members of the Borrelia genus

XX Sequence 290 AA;

Query Match 99.6%; Score 468; DB 2; Length 290;

Best Local Similarity 98.9%; Pred. No. 2.9e-49;

Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQGFVVPSPVPISSISELKGKDKFKNMIGIDAGAGTQIVTEQALNYGLSKYELVPS 60

Db 109 IQGFVVPSPVPISSISELKGKDKFKNMIGIDAGAGTQIVTEQALNYGLSKYELVPS 168

QY 61 SESVMLASLDSSIKRNEWILVPLMKPHWAFS 91

Db 169 SESVMLASLDSSIKRNEWILVPLMKPHWAFS 199

PT Enterococcus faecium polypeptide useful for detection, prevention and
 PT treatment of a pathological condition resulting from a bacterial
 PT infection.
 XX
 XX
 PS Example 1; SEQ ID NO 5217; 243pp; English.
 XX
 XX The invention relates to an isolated nucleic acid derived from
 CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
 CC one of 10 fully defined sequences given in the (or comprising 40
 CC sequential nucleotides chosen from any of the nucleic acids, its
 CC complement or sequences hybridising to it). Also included are a
 CC recombinant vector comprising the nucleic acid operably linked to
 CC transcription regulatory element, a cell comprising the vector and a
 CC single-stranded probe comprising the nucleic acid. The nucleic acids are
 CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
 CC The nucleic acids are useful for diagnosing pathological conditions
 CC resulting from E. faecium bacterial infection (e.g. urinary tract
 CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
 CC infection) and for screening drugs such as agonists and antagonists. The
 CC nucleic acid is useful for recombinant production of Candida albicans -
 CC derived peptides or antisense polypeptides. Pharmaceutical compositions
 CC and vaccines containing the nucleic acid are useful for preventing or
 CC treating Enterococcus faecium infections. The present sequence represents
 CC one if the disclosed E. faecium proteins.
 XX

Sequence 577 AA;

Query Match 34.4%; Score 161.5; DB 7; Length 577;
 Best Local Similarity 37.1%; Pred. No. 9.6e-11;
 Matches 33; Conservative 18; Mismatches 35; Indels 3; Gaps 1;

QY 3 GFVPSYVPISSISLKGKDKFKNMIGIDAGAGTQIVTEQALNYGLSKYEYLVPSSE 62
 DB 410 GLAYPAYDVNSIDELTQAGK--KIIGIEPGAGVVTAAENTIOKYNLKDQKVTETSS 466

QY 63 SVMLASLDSIKRNEWILVPLWKPHWAFS 91

DB 467 GANTVALGQAIIKKEPIVVTGWTWHMFA 495

RESULT 9

AAU34910
 ID AAU34910 standard; protein; 569 AA.

XX AAU34910;

XX 14-FEB-2002 (first entry)

XX Enterococcus faecalis cellular proliferation protein #197.

XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW- antibacterial; drug design.

OS Enterococcus faecalis.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2000; 2000US-0191078P.

PR 23-MAY-2000; 2000US-0206848P.

PR 26-MAY-2000; 2000US-0207727P.

PR 23-OCT-2000; 2000US-0242578P.

PR 27-NOV-2000; 2000US-0253625P.

PR 22-DEC-2000; 2000US-0257931P.

PR 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.
 DR N-PSDB; AAS52769.

XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.

XX Example 3; SEQ ID NO 10503; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes,
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 569 AA;

Query Match 33.7%; Score 158.5; DB 4; Length 569;
 Best Local Similarity 37.1%; Pred. No. 2.2e-10;
 Matches 33; Conservative 15; Mismatches 38; Indels 3; Gaps 1;

QY 3 GFVPSYVPISSISLKGKDKFKNMIGIDAGAGTQIVTEQALNYGLSKYEYLVPSSE 62

DB 402 GIVPFSYMDVDSIEDLSQAGK--KITGIEPGAGVAAAEKTKAYPNLKDQSVETSS 458

QY 63 SVMLASLDSIKRNEWILVPLWKPHWAFS 91

DB 459 GANTVALGQAIIKKNEDIVITGWSPHMFA 487

RESULT 10

ADH87483

ID ADH87483 standard; protein; 571 AA.

XX ADH87483;

XX 22-APR-2004 (first entry)

XX Enterococcus faecalis polypeptide #1963.

XX Enterococcus faecalis infection; transcription regulatory element;
 KW antibacterial.

OS Enterococcus faecalis.

XX US6617156-B1.

XX 09-SEP-2003.

XX 13-AUG-1998; 98US-00134000.

XX 15-AUG-1997; 97US-0055778P.

XX (DOUC/) DOUCETTE-STAMM L A.

PA (BUSH/) BUSH D.

XX Doucette-Stamm LA, Bush D;

XX WPI; 2003-895394/82.

DR N-PSDB; ADH84078.

XX New nucleic acid comprising a sequence encoding an Enterococcus faecalis
PT polypeptide, useful for preparing a composition for diagnosing or
PT treating E. faecalis infection.
XX
PS Disclosure; SEQ ID NO 5368; 193pp; English.
XX
CC The invention relates to Enterococcus faecalis polynucleotides and
CC polypeptides. The invention also relates to a recombinant expression
CC vector comprising a polynucleotide operably linked to a transcription
CC regulatory element, a cell comprising a recombinant vector, a method for
CC producing an E. faecalis polypeptide, an isolated nucleic acid comprising
CC a sequence not given in the specification, a recombinant vector
CC comprising the nucleic acid and a cell comprising the recombinant vector.
CC The polynucleotides can be used to detect the presence of E. faecalis in
CC a sample. The sequences are useful for preparing a composition for
CC diagnosing or treating Enterococcus faecalis infection. This sequence
CC represents an E. faecalis polypeptide of the invention.
XX
SQ Sequence 571 AA;

Query Match 33.7%; Score 158.5; DB 7; Length 571;
Best Local Similarity 37.1%; Pred. No. 2.2e-10;
Matches 33; Conservative 15; Mismatches 38; Indels 3; Gaps 1;
QY 3 GFVPSYVPISSISELKGKDKFKNMIGIDAGAGTQIVTEQALNYGLSKEYELVPSSE 62
DB 404 GIVPSYMDVSDIEDLSDAQK---KITGIEFGVVAANAETKEAYPNLKDMSVETSSS 460
QY 63 SYMLASLDSSIKRNEWILVPLWKPWFAS 91
DB 461 GANTVALGQAIKNNEDIVTGSFWMFA 489

RESULT 11
ID ABP27731 standard; protein; 575 AA.
XX
AC ABP27731;
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 4638.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus pyogenes.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB004789.
XX
PR 27-OCT-2000; 2000GB-00026333.
PR 24-NOV-2000; 2000GB-00028727.
PR 07-MAR-2001; 2001GB-00005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Masignani V, Margarit Y Rosl, Grandi G, Fraser C;
PI Tettelin H;
XX
DR WPI; 2002-352536/38.
DR N-PSDB; ABN68362.
XX
PT New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.
XX

PS Claim 1; Page 3626-3627; 4525pp; English.
XX
CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX
SQ Sequence 575 AA;

Query Match 33.4%; Score 157; DB 5; Length 575;
Best Local Similarity 36.7%; Pred. No. 3.5e-10;
Matches 33; Conservative 17; Mismatches 36; Indels 4; Gaps 2;
QY 3 GFVPSYVPISSISELKGKDKFKNMIGIDAGAGTQIVTEQALNYGLSKEYELVPSSE 61
DB 407 GLAVPKYMTDVSIEDLSKQADQ---KITGIEFGAGIMAAQKTUKEYHNLSWELVAAS 463
QY 62 ESMVLASLDSSIKRNEWILVPLWKPWFAS 91
DB 464 TGMATTSILDQAIAKKDPIVVTANSPHWFA 493

RESULT 12
ID ADB07854 standard; protein; 273 AA.
XX
AC ADB07854;
XX
DT 20-NOV-2003 (first entry)
XX
DE Alloicoccus otitis antigenic protein SEQ ID NO:1794.
XX
KW Alloicoccus otitis; antigenic protein; immunogenic; immunisation;
KW gene therapy; Gram-positive bacterium; infection.
XX
OS Alloicoccus otitis.
XX
PN WO2003048304-A2.
XX
PD 12-JUN-2003.
XX
PF 25-NOV-2002; 2002WO-US036123.
XX
PR 29-NOV-2001; 2001US-0333777P.
PR 18-NOV-2002; 2002US-0426742P.
XX
PA (AMHP) WYETH HOLDINGS CORP.
XX
PI Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;
XX
DR WPI; 2003-505284/47.
DR N-PSDB; ADB07853.
XX
PT New Alloicoccus otitis polynucleotides and polypeptides, useful for
PT treating and diagnosing diseases, drug screening assays and monitoring of
PT effects during drug clinical trials.
XX
PS Claim 33; SEQ ID NO 1794; 1019pp; English.
XX
CC The present invention describes an isolated polynucleotide (I) of


```
PI Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;
XX WPI; 2003-505284/47.
DR N-PSDB; ADB07857.
XX
XX New Alloiooccus otitidis polynucleotides and polypeptides, useful for
PT treating and diagnosing diseases, drug screening assays and monitoring of
PT effects during drug clinical trials.
XX
XX Claim 33; SEQ ID NO 1798; 1019pp; English.
XX
XX The present invention describes an isolated polynucleotide (1) of
CC Alloiooccus otitidis genomic DNA, which encodes an antigenic protein.
CC Alloiooccus otitidis is a Gram-positive bacterium. Also described: (1)
CC an isolated polypeptide that is encoded by the polynucleotide (1); (2) an
CC expression vector comprising the novel isolated polynucleotide (1), its
CC complement, degenerate variant or fragment; (3) a genetically engineered
CC host cell, transfected, transformed or infected with the vector of (2);
CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
CC composition comprising the polypeptide, its complement, biological
CC equivalent or fragment, or the polynucleotide that is comprised in the
CC expression vector; (6) a pharmaceutical composition comprising the
CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
CC of the polypeptides of (1), their biological equivalent or fragment; (8)
CC immunising against Alloiooccus otitidis by administering to a host the
CC immunogenic composition; (9) detecting and/or identifying Alloiooccus
CC otitidis in the biological sample; (10) a kit comprising a container
CC containing the novel polynucleotide, its degenerate variant or fragment,
CC or the antibody of (4); and (11) producing a polypeptide by culturing the
CC genetically engineered host cell under conditions suitable to produce the
CC polypeptide from the culture. (1) can be used in gene therapy. The
CC polynucleotides, polypeptides, antibodies and compositions of the present
CC invention can be used for treating and diagnosing diseases, drug
CC screening assays and monitoring of effects during drug clinical trials.
CC The polynucleotides are useful for expressing and detecting Alloiooccus
CC otitidis. The present sequence represents an Alloiooccus otitidis
CC antigen protein from the present invention.
XX
XX Sequence 352 AA;
XX
XX Query Match 32.6%; Score 153; DB 6; Length 352;
XX Best Local Similarity 39.1%; Pred. No. 5.6e-10;
XX Matches 36; Conservative 19; Mismatches 31; Indels 6; Gaps 4;
XX
XX QY 1 LOGFVVPVSVVP-SSISELKGKDKFNK-MIGIDAGAGTQIVTQALNYGLSKEYEIV 58
XX 123 LGLVVPYMEGDSIEDUE---DQEDKTIYGIERGAGTVEAERALEDDYG-KEDWEVI 178
XX
XX QY 59 PSESVMLASLDSSIKRNEWILVPLWKPHWAF 90
XX 179 ASSSGAMVTELRAQINNEEDIIVTGTWPHKF 210
XX
XX Db
XX
XX RESULT 15
XX ADB07852
XX ID ADB07852 standard; protein; 320 AA.
XX AC ADB07852;
XX
XX XX 20-NOV-2003 (first entry)
XX
XX DE Alloiooccus otitis antigenic protein SEQ ID NO:1792.
XX
XX KW Alloiooccus otitidis; antigenic protein; immunogenic; immunisation;
XX gene therapy; Gram-positive bacterium; infection.
XX
XX OS Alloiooccus otitis.
XX
XX PN WO2003048304-A2.
XX
XX PD 12-JUN-2003.
XX
XX PF 25-NOV-2002; 2002WO-US036123.
XX
```

```
XX
XX 29-NOV-2001; 2001US-0333777P.
XX 18-NOV-2002; 2002US-0426742P.
XX
XX (AMHP ) WYETH HOLDINGS CORP.
XX
XX Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;
XX
XX WPI; 2003-505284/47.
XX N-PSDB; ADB07851.
XX
XX New Alloiooccus otitidis polynucleotides and polypeptides, useful for
PT treating and diagnosing diseases, drug screening assays and monitoring of
PT effects during drug clinical trials.
XX
XX Claim 33; SEQ ID NO 1792; 1019pp; English.
XX
XX The present invention describes an isolated polynucleotide (1) of
CC Alloiooccus otitidis genomic DNA, which encodes an antigenic protein.
CC Alloiooccus otitidis is a Gram-positive bacterium. Also described: (1)
CC an isolated polypeptide that is encoded by the polynucleotide (1); (2) an
CC expression vector comprising the novel isolated polynucleotide (1), its
CC complement, degenerate variant or fragment; (3) a genetically engineered
CC host cell, transfected, transformed or infected with the vector of (2);
CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
CC composition comprising the polypeptide, its complement, biological
CC equivalent or fragment, or the polynucleotide that is comprised in the
CC expression vector; (6) a pharmaceutical composition comprising the
CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
CC of the polypeptides of (1), their biological equivalent or fragment; (8)
CC immunising against Alloiooccus otitidis by administering to a host the
CC immunogenic composition; (9) detecting and/or identifying Alloiooccus
CC otitidis in the biological sample; (10) a kit comprising a container
CC containing the novel polynucleotide, its degenerate variant or fragment,
CC or the antibody of (4); and (11) producing a polypeptide by culturing the
CC genetically engineered host cell under conditions suitable to produce the
CC polypeptide from the culture. (1) can be used in gene therapy. The
CC polynucleotides, polypeptides, antibodies and compositions of the present
CC invention can be used for treating and diagnosing diseases, drug
CC screening assays and monitoring of effects during drug clinical trials.
CC The polynucleotides are useful for expressing and detecting Alloiooccus
CC otitidis. The present sequence represents an Alloiooccus otitidis
CC antigen protein from the present invention.
XX
XX Sequence 320 AA;
XX
XX Query Match 32.4%; Score 152.5; DB 6; Length 320;
XX Best Local Similarity 36.4%; Pred. No. 5.7e-10;
XX Matches 32; Conservative 17; Mismatches 36; Indels 3; Gaps 1;
XX
XX QY 3 GFVVPVYVPISSISELKGKDKFNKMGIDAGAGTQIVTQALNYGLSKEYELVPSSE 62
XX 122 GLGVPTMEVDSIADLT---DEAEQTITGIEAGAGVQSGAEQALDDYDNLSDWTATSSS 178
XX
XX Db
XX
XX QY 63 SVMLASLDSSIKRNEWILVPLWKPHWAF 90
XX 179 GAMVTELETAIINNEEDIIVTAWSPHKF 206
XX
XX Db
XX
XX Search completed: January 24, 2006, 19:44:28
XX Job time : 22.3361 secs
XX
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:30:59 ; Search time 5.60996 Seconds
(without alignments)
1341.095 Million cell updates/sec

Title: US-10-688-058-6
Perfect score: 470
Sequence: 1 LQGFVVPVPISSISLKGK.....SIKRNWILVPLWKPWFAPS 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/1/1aa/5 COMB.pep.*
 - 2: /cgn2_6/ptodata/1/1aa/6 COMB.pep.*
 - 3: /cgn2_6/ptodata/1/1aa/H COMB.pep.*
 - 4: /cgn2_6/ptodata/1/1aa/PCTUS COMB.pep.*
 - 5: /cgn2_6/ptodata/1/1aa/RE COMB.pep.*
 - 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	468	99.6	275	2	US-09-830-230A-362
2	468	99.6	290	2	US-09-830-230A-361
3	193	41.1	384	2	US-09-252-991A-21729
4	179	38.1	309	2	US-09-902-540-15605
5	161.5	34.4	577	2	US-09-107-532A-5217
6	158.5	33.7	571	2	US-09-134-000C-5368
7	138	29.4	318	2	US-09-252-991A-17173
8	127	27.0	356	2	US-09-252-991A-17157
9	126	26.8	427	2	US-09-252-991A-17157
10	88	18.7	317	2	US-09-252-991A-31329
11	86	18.3	338	2	US-09-489-039A-10172
12	76	16.2	361	2	US-09-252-991A-32937
13	74	15.7	573	2	US-09-134-001C-4656
14	69	14.7	426	2	US-09-807-258-26
15	68	14.5	317	2	US-09-489-039A-9963
16	67.5	14.4	1208	2	US-09-540-236-3604
17	67.5	14.4	1218	2	US-09-198-452A-98
18	67.5	14.4	1410	2	US-09-438-185A-84
19	67	14.3	289	2	US-09-807-258-22
20	67	14.3	379	2	US-09-248-796A-18875
21	66	14.0	251	2	US-09-107-532A-5242
22	66	14.0	617	2	US-09-565-264-2
23	66	14.0	617	2	US-09-565-264-4
24	66	14.0	646	2	US-09-252-991A-16819
25	66	14.0	1059	2	US-09-134-000C-5600
26	65.5	13.9	364	2	US-08-818-112-73
27	65.5	13.9	364	2	US-08-818-111-74

28	65.5	13.9	364	2	US-09-056-556-73	Sequence 73, Appl
29	65.5	13.9	364	2	US-09-072-596-74	Sequence 74, Appl
30	65.5	13.9	364	2	US-09-072-967-73	Sequence 73, Appl
31	65.5	13.9	364	2	US-10-193-002-74	Sequence 74, Appl
32	65.5	13.9	364	2	US-10-084-843-73	Sequence 73, Appl
33	65	13.8	1215	2	US-09-134-001C-5319	Sequence 5319, Ap
34	64	13.6	442	2	US-08-821-994-66	Sequence 66, Appl
35	63.5	13.5	837	2	US-09-122-126B-2	Sequence 2, Appl
36	63.5	13.5	837	2	US-09-634-286A-2	Sequence 2, Appl
37	63.5	13.5	837	2	US-10-247-685-2	Sequence 2, Appl
38	63.5	13.5	837	2	US-09-949-002-380	Sequence 380, App
39	63.5	13.5	837	2	US-10-012-231A-317	Sequence 317, App
40	63.5	13.5	837	2	US-10-015-389A-317	Sequence 317, App
41	63.5	13.5	837	2	US-10-006-768A-317	Sequence 317, App
42	63.5	13.5	837	2	US-10-015-671A-317	Sequence 317, App
43	63.5	13.5	837	2	US-10-015-393A-317	Sequence 317, App
44	63.5	13.5	837	2	US-10-011-833A-317	Sequence 317, App
45	63.5	13.5	837	2	US-10-006-041A-317	Sequence 317, App

ALIGNMENTS

RESULT 1

US-09-830-230A-362
; Sequence 362, Application US/09830230A
; Patent No. 6902893
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481US
; CURRENT APPLICATION NUMBER: US/09/830,230A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: PCT/US98/12718
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/057,483
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 362
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-830-230A-362

Query Match 99.6%; Score 468; DB 2; Length 275;
Best Local Similarity 98.9%; Pred. No. 3e-52;
Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQGFVVPVPISSISLKGKDKFKNMIGIDAGAGTQIVTEQALNYVGLSKEYELVPS 60
DB 94 IQGFVVPVPISSISLKGKDKFKNMIGIDAGAGTQIVTEQALNYVGLSKEYELVPS 153
QY 61 SESVNLASLDSISKKNWILVPLWKPWFAPS 91
DB 154 SESVNLASLDSISKKNWILVPLWKPWFAPS 184

RESULT 2

US-09-830-230A-361
; Sequence 361, Application US/09830230A
; Patent No. 6902893
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481US
; CURRENT APPLICATION NUMBER: US/09/830,230A

```
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: PCT/US98/12718
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/057,483
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 361
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-830-230A-361

Query Match          99.6%; Score 468; DB 2; Length 290;
Best Local Similarity 98.9%; Pred. No. 3.2e-52;
Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQGFVVPSTYPISSISLKGKGDGKFKNMIGIDAGAGTQIVTEQALNYYGLSKYELVPS 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 109 LQGFVVPSTYPISSISLKGKGDGKFKNMIGIDAGAGTQIVTEQALNYYGLSKYELVPS 168

QY 61 SESVMLASLDSSIKRNEWILVPLWKPHWAFS 91
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 169 SESVMLASLDSSIKRNEWILVPLWKPHWAFS 199
```

```
RESULT 3
US-09-252-991A-21729
; Sequence 21729, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21729
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21729
```

```
Query Match          41.1%; Score 193; DB 2; Length 384;
Best Local Similarity 42.7%; Pred. No. 2.1e-16;
Matches 38; Conservative 20; Mismatches 29; Indels 2; Gaps 1;

QY 3 GFVVPSTYPISSISLKGKGDGKFKNMIGIDAGAGTQIVTEQALNYYGLSKYELVPS 62
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 215 GLIVPEYVSANSTADLAQKDAFGGRVVGIDAGAGVMIKTDEAIKQYGL--DYKLVASG 272

QY 63 SVMLASLDSSIKRNEWILVPLWKPHWAFS 91
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 273 SGMIAEIATRAENEKPPVVVTGIPHMWFA 301
```

```
RESULT 4
US-09-902-540-15605
; Sequence 15605, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
```

```
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15605
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15605

Query Match          38.1%; Score 179; DB 2; Length 309;
Best Local Similarity 41.6%; Pred. No. 1e-14;
Matches 37; Conservative 16; Mismatches 34; Indels 2; Gaps 1;

QY 3 GFVVPSTYPISSISLKGKGDGKFKNMIGIDAGAGTQIVTEQALNYYGLSKYELVPS 62
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 138 GLVVPADLDTISIQLNKAKDLNQSIVGIDSGAGIMTTTTEKATAYKL--DLKLVPSG 195

QY 63 SVMLASLDSSIKRNEWILVPLWKPHWAFS 91
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 196 PAMTAEIKDAIAKRPVVVTGKPHWKEA 224

RESULT 5
US-09-107-532A-5217
; Sequence 5217, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5217:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 577 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
```

```
; ORGANISM: Enterococcus faecium
;
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...577
; SEQUENCE DESCRIPTION: SEQ ID NO: 5217:
US-09-107-532A-5217

Query Match      34.4%; Score 161.5; DB 2; Length 577;
Best Local Similarity 37.1%; Pred. No. 4.5e-12;
Matches 33; Conservative 18; Mismatches 35; Indels 3; Gaps 1;

QY      3 GFVPSYVPISSELKGGDKFKNMIGIDAGAGTQIVTEQALNYGLSKEYELVPSSE 62
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db      410 GLAVPAYMDVNSIDELTQAGK---KIIEFGAGVTTAAENTIOKYNLKDWSVETSSS 466
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

QY      63 SVMLASLDSSIKRNEWILVPLWKPHWAFS 91
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db      467 GANTVALGQAIIKKEPIVVTGTPHMFPA 495
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

RESULT 6
US-09-134-000C-5368
; Sequence 5368, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5368
; LENGTH: 571
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5368

Query Match      33.7%; Score 158.5; DB 2; Length 571;
Best Local Similarity 37.1%; Pred. No. 1.1e-11;
Matches 33; Conservative 15; Mismatches 38; Indels 3; Gaps 1;

QY      3 GFVPSYVPISSELKGGDKFKNMIGIDAGAGTQIVTEQALNYGLSKEYELVPSSE 62
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db      404 GIVPSYMDVDSIEDLSQAGK---KITGIEFGAGVAAAEKTKAYPNLKDWSVETSSS 460
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

QY      63 SVMLASLDSSIKRNEWILVPLWKPHWAFS 91
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db      461 GANTVALGQAIIKNNEDIVITGWSPHMFPA 489
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

RESULT 7
US-09-252-991A-17173
; Sequence 17173, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17173
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17173
```

```
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17173

Query Match      29.4%; Score 138; DB 2; Length 318;
Best Local Similarity 33.7%; Pred. No. 2.2e-09;
Matches 30; Conservative 24; Mismatches 29; Indels 6; Gaps 3;

QY      4 FVPSYVP---ISSISELKGDKFKNMIGIDAGAGTQIVTEQAL--NYGLSKEYELV 58
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db      123 YAVPSYVAEGKTFADIARFKDKLGGRYIAIEPGSGSNRITRKWIDNRFGL-KGFQLV 181
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

QY      59 PSSSEVMSLASLDSSIKRNEWILVPLWKPH 87
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db      182 ESSERAGMLTAVKRAIKRQWVFFGWKPH 210
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

RESULT 8
US-09-252-991A-17157
; Sequence 17157, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17157
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17157

Query Match      27.0%; Score 127; DB 2; Length 356;
Best Local Similarity 34.5%; Pred. No. 7e-08;
Matches 30; Conservative 20; Mismatches 31; Indels 6; Gaps 3;

QY      6 VPSYV---PISISELKGDKFKNMIGIDAGAGTQIVTEQAL--NYGLSKEYELVPS 60
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db      161 VPQVYDGDLSFADIAKFSKLGKNIYGIIEPGDGNRVAQSMIDKNFELGK-FKLVES 219
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

QY      61 SESVMSLASLDSSIKRNEWILVPLWKPH 87
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db      220 SEAGMLSOVQAIRRNQWVFLGWEPH 246
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

RESULT 9
US-09-252-991A-21769
; Sequence 21769, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21769
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21769
```


QY 64 VMLASLDSISIKRNE 77
Db 500 QGLAWMLSYVK-NE 512

RESULT 14

US-09-807-258-26
; Sequence 26, Application US/09807258
; Patent No. 6670186
; GENERAL INFORMATION:
; APPLICANT: E. I. du Pont de Nemours and Company
; TITLE OF INVENTION: Arthropod Protein Disulfide Isomerases
; FILE REFERENCE: BB-1253 PCT
; CURRENT APPLICATION NUMBER: US/09/807,258
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/104,376
; PRIOR FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 26
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Spodoptera frugiperda
US-09-807-258-26

Query Match 14.7%; Score 69; DB 2; Length 426;
Best Local Similarity 32.3%; Pred. No. 3;
Matches 20; Conservative 13; Mismatches 25; Indels 4; Gaps 2;
QY 22 GDKPKNKMG-IDAGAGTQIVTEQALNYGLSKYELVPSSVMLASLDSIKR--NE 77
Db 313 GDKYKKNKGWVWAEAGAGLGLSELELGGFGYPAMAVVNAKXKLFSTLRGSFSETGINE 372

QY 78 WI 79
Db 373 FL 374

RESULT 15

US-09-489-039A-9963
; Sequence 9963, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709-2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9963
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9963

Query Match 14.5%; Score 68; DB 2; Length 317;
Best Local Similarity 26.9%; Pred. No. 2.7;
Matches 25; Conservative 18; Mismatches 30; Indels 20; Gaps 6;
QY 2 QGFWVPSPVPISSISLKGKDGKPKNKMIGIDAGAGTQIVTEQALNYGLSKYELVPSS 61
Db 115 EAIVVPEQSATHSVADLKG-----KRVGLNKGSDVNVLLVALEKAGLSYK-DITP-- 164
QY 62 ESMVLASLD--SSIKR---NEWILVPLMKPHWA 89
Db 165 --VYLPADARAFAFGAIDAWV---IWDPFLLA 192

Search completed: January 24, 2006, 20:02:39
Job time : 5.60996 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:57:15 ; Search time 15.9338 Seconds
(without alignments)
2377.329 Million cell updates/sec

Title: US-10-688-058-6

Perfect score: 470

Sequence: 1 LQGFVPSVYPISSISLKG.....SIKRNWILVPLMKPHWAFS 91

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries .

Database :

Published Applications AA Main:
1: /cgm2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgm2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgm2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgm2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
5: /cgm2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgm2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	470	100.0	91	5	US-10-688-058-6
2	468	99.6	275	5	US-10-994-726-362
3	468	99.6	290	5	US-10-688-058-8
4	468	99.6	290	5	US-10-994-726-361
5	192	40.9	871	4	US-10-156-761-14244
6	158.5	33.7	569	3	US-09-815-242-10503
7	157	33.4	575	5	US-10-474-792-100
8	153	32.6	273	5	US-10-501-282-1794
9	153	32.6	334	5	US-10-501-282-1796
10	153	32.6	352	5	US-10-501-282-1798
11	152.5	32.4	320	5	US-10-501-282-1792
12	138	29.4	312	3	US-09-815-242-5213
13	138	29.4	312	4	US-10-282-122A-43548
14	130	27.7	302	5	US-10-501-282-5510
15	127	27.0	380	4	US-10-282-122A-49743
16	124	26.4	303	5	US-10-501-282-1918
17	124	26.4	326	5	US-10-501-282-1920
18	108	23.0	314	4	US-10-282-122A-67487
19	87	18.5	332	4	US-10-156-761-10018
20	76.5	16.3	903	4	US-10-282-122A-77450
21	74	15.7	319	4	US-10-032-201B-200
22	74	15.7	573	4	US-10-724-972A-5090
23	71.5	15.2	773	4	US-10-437-963-162371
24	70	14.9	314	4	US-10-032-201B-201
25	69.5	14.8	581	4	US-10-437-963-122382
26	69.5	14.8	839	3	US-09-912-788-2
27	69.5	14.8	1178	4	US-10-282-122A-52187

28 69 14.7 1161 4 US-10-282-122A-52986 Sequence 52986, A
29 68.5 14.6 869 5 US-10-732-923-6769 Sequence 6769, Ap
30 68.5 14.6 1106 5 US-10-501-282-466 Sequence 466, App
31 68.5 14.6 1128 5 US-10-501-282-468 Sequence 468, App
32 68.5 14.6 1192 5 US-10-501-282-470 Sequence 470, App
33 68.5 14.6 1219 5 US-10-501-282-472 Sequence 472, App
34 68 14.5 299 4 US-10-282-122A-60579 Sequence 60579, A
35 68 14.5 424 4 US-10-282-122A-48661 Sequence 48661, A
36 68 14.5 461 5 US-10-482-076-6 Sequence 271, App
37 68 14.5 461 5 US-10-482-076-271 Sequence 51577, A
38 67.5 14.4 1182 4 US-10-282-122A-51577 Sequence 51577, A
39 67.5 14.4 1203 4 US-10-282-122A-45530 Sequence 45530, A
40 67.5 14.4 1218 4 US-10-289-762-98 Sequence 98, Appl
41 67.5 14.4 1396 4 US-10-282-122A-55213 Sequence 55213, A
42 67.5 14.4 1397 4 US-10-282-122A-54874 Sequence 54874, A
43 67.5 14.4 2560 4 US-10-276-774-1774 Sequence 1774, Ap
44 67.5 14.4 2923 3 US-09-788-711A-4 Sequence 4, Appli
45 67.5 14.4 2923 3 US-09-916-849A-3 Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-10-688-058-6
; Sequence 6, Application US/10688058
; Publication No. US20050058661A1
; GENERAL INFORMATION:
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: HALE, KATHERINE S.
; APPLICANT: JOHNSTON, STEPHEN A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF THE GENUS
; TITLE OF INVENTION: BORRELIA
; FILE REFERENCE: MCRO:00305
; CURRENT APPLICATION NUMBER: US/10/688,058
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-688-058-6

Query Match 100.0%; Score 470; DB 5; Length 91;
Best Local Similarity 100.0%; Pred. No. 3e-48;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LQGFVPSVYPISSISLKGDKFKFKMIGIDAGAGTQIVTEQALNYGLSKVELVPS 60
DB 1 LQGFVPSVYPISSISLKGDKFKFKMIGIDAGAGTQIVTEQALNYGLSKVELVPS 60
QY 61 SESVNLASLDSIKRNEMILVPLMKPHWAFS 91
DB 61 SESVNLASLDSIKRNEMILVPLMKPHWAFS 91
RESULT 2
US-10-994-726-362
; Sequence 362, Application US/10994726
; Publication No. US20050147999A1
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481D1
; CURRENT APPLICATION NUMBER: US/10/994,726
; CURRENT FILING DATE: 2004-11-23
; PRIOR APPLICATION NUMBER: 09/830,230
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: PCT/US98/12718
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/057,483

```
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 362
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-994-726-362

Query Match          99.6%; Score 468; DB 5; Length 275;
Best Local Similarity 98.9%; Pred. No. 2.1e-47;
Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQGFVVPSPVPISSISELKGDKFKNKMIGIDAGAGTQIVTEQALNYGLSKYELVPS 60
Db 94 IQGFVVPSPVPISSISELKGDKFKNKMIGIDAGAGTQIVTEQALNYGLSKYELVPS 153
QY 61 SESVMLASLDSSIKRNEWILVPLWKPHWAFS 91
Db 154 SESVMLASLDSSIKRNEWILVPLWKPHWAFS 184

RESULT 3
US-10-688-058-8
; Sequence 8, Application US/10688058
; Publication No. US20050058661A1
; GENERAL INFORMATION:
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: HALE, KATHERINE S.
; APPLICANT: JOHNSTON, STEPHEN A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF THE GENUS
; TITLE OF INVENTION: BORRELLIA
; FILE REFERENCE: MCRO:003US
; CURRENT APPLICATION NUMBER: US/10/688,058
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-688-058-8

Query Match          99.6%; Score 468; DB 5; Length 290;
Best Local Similarity 98.9%; Pred. No. 2.3e-47;
Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQGFVVPSPVPISSISELKGDKFKNKMIGIDAGAGTQIVTEQALNYGLSKYELVPS 60
Db 109 IQGFVVPSPVPISSISELKGDKFKNKMIGIDAGAGTQIVTEQALNYGLSKYELVPS 168
QY 61 SESVMLASLDSSIKRNEWILVPLWKPHWAFS 91
Db 169 SESVMLASLDSSIKRNEWILVPLWKPHWAFS 199

RESULT 4
US-10-994-726-361
; Sequence 361, Application US/10994726
; Publication No. US20050147999A1
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481D1
; CURRENT APPLICATION NUMBER: US/10/994,726
; CURRENT FILING DATE: 2004-11-23
```

```
; PRIOR APPLICATION NUMBER: 09/830,230
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: PCT/US98/12718
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/057,483
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 361
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-994-726-361

Query Match          99.6%; Score 468; DB 5; Length 290;
Best Local Similarity 98.9%; Pred. No. 2.3e-47;
Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQGFVVPSPVPISSISELKGDKFKNKMIGIDAGAGTQIVTEQALNYGLSKYELVPS 60
Db 109 IQGFVVPSPVPISSISELKGDKFKNKMIGIDAGAGTQIVTEQALNYGLSKYELVPS 168
QY 61 SESVMLASLDSSIKRNEWILVPLWKPHWAFS 91
Db 169 SESVMLASLDSSIKRNEWILVPLWKPHWAFS 199

RESULT 5
US-10-156-761-14244
; Sequence 14244, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14244
; LENGTH: 871
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14244

Query Match          40.9%; Score 192; DB 4; Length 871;
Best Local Similarity 45.5%; Pred. No. 1.2e-13;
Matches 40; Conservative 15; Mismatches 31; Indels 2; Gaps 2;

QY 6 VPSYV-PISSISELKGDKFKNKMIGIDAGAG-TQIVTEQALNYGLSKYELVPSSES 63
Db 699 VPSYVKDVKSLADLKGKGTEDGKIIGIEPTGEMNLLKTKVLPGYGLDKEYKVVDGSGTP 758
QY 64 VMLASLDSSIKRNEWILVPLWKPHWAFS 91
Db 759 AMLAEKRAYAKKEPVAVVLMSPHWAFS 786
```

RESULT 6

US-09-815-242-10503

; Sequence 10503, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10503

; LENGTH: 569

; TYPE: PRT

; ORGANISM: Enterococcus faecalis

US-09-815-242-10503

Query Match

Best Local Similarity 33.7%; Score 158.5; DB 3; Length 569;

Matches 33; Conservative 15; Mismatches 38; Indels 3; Gaps 1;

QY 3 GFVVPVSVPISSISLKGKDGKFKNMIGIDAGAGTQIVTEQALNYGLSKEYELVPSSE 62

DB 402 GIVPVSMDVSDIEDLSQAGK---KITGIEPGAGVAAAEKTKAAYPNLKDWSVETSSS 458

QY 63 SVMLASLDSSIKRNEWILVPLWKPWFAPS 91

DB 459 GANTVALGQAIRKNNEDIVTGNWSPHMFA 487

RESULT 7

US-10-474-792-100

; Sequence 100, Application US/10474792

; Publication No. US20040236072A1

; GENERAL INFORMATION:

; APPLICANT: Olmsted, Stephen

; APPLICANT: Zagursky, Robert

; APPLICANT: Nickbarg, Elliot

; APPLICANT: Winter, Lourie

; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES

; FILE REFERENCE: AM 100399

; CURRENT APPLICATION NUMBER: US/10/474,792

; CURRENT FILING DATE: 2003-10-14

; NUMBER OF SEQ ID NOS: 674

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 100

; LENGTH: 575

; TYPE: PRT

; ORGANISM: Streptococcus pyogenes

US-10-474-792-100

Query Match

Best Local Similarity 33.4%; Score 157; DB 5; Length 575;

Matches 33; Conservative 17; Mismatches 36; Indels 4; Gaps 2;

QY 3 GFVVPVSVPISSISLKGKDGKFKNMIGIDAGAGTQIVTEQALNYGLSKEYELVPSSE 61

DB 407 GLAVPKYMTDVNSIEDLSKQADQ---KITGIEPGAGVAAAKTKLKEYHNLSSWELVAAS 463

QY 62 ESMVLASLDSSIKRNEWILVPLWKPWFAPS 91

DB 464 TGAMTTSILDQAIRKNNEDIVTGNWSPHMFA 493

RESULT 8

US-10-501-282-1794

; Sequence 1794, Application US/10501282

; Publication No. US20050203280A1

; GENERAL INFORMATION:

; APPLICANT: MCMICHAEL, JOHN CALHOUN

; APPLICANT: ZAGURSKY, ROBERT JOHN

; APPLICANT: RUSSELL, DAVID PARRISH

; APPLICANT: FLETCHER, LEAH DIANE

; TITLE OF INVENTION: ALLOIOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING

; TITLE OF INVENTION: POLYPEPTIDE ANTIGENS, IMMUNOGENIC COMPOSITIONS AND USES THEREOF

; FILE REFERENCE: AM100780 L2

; CURRENT APPLICATION NUMBER: US/10/501,282

; CURRENT FILING DATE: 2004-07-09

; PRIOR APPLICATION NUMBER: 60/333,777

; PRIOR FILING DATE: 2001-11-29

; PRIOR APPLICATION NUMBER: 60/426,742

; PRIOR FILING DATE: 2002-11-18

; PRIOR APPLICATION NUMBER: PCT/US02/36123

; PRIOR FILING DATE: 2002-11-25

; NUMBER OF SEQ ID NOS: 6653

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1794

; LENGTH: 273

; TYPE: PRT

; ORGANISM: Alloiococcus otitidis

US-10-501-282-1794

Query Match

Best Local Similarity 32.6%; Score 153; DB 5; Length 273;

Matches 36; Conservative 19; Mismatches 31; Indels 6; Gaps 4;

QY 1 LOGFVVPVSVPISSISLKGKDGKFKNMIGIDAGAGTQIVTEQALNYGLSKEYELV 58

DB 44 LIGLVVPYMEGLDSIEDLE---DOIETKIIVGIERGAGTVEAERALEDYG-KEDWEVI 99

QY 59 PSSESVMLASLDSSIKRNEWILVPLWKPWFAPS 90

DB 100 ASSGAMVTELRQAINNEEDIVTGTWTHKPF 131

RESULT 9

US-10-501-282-1796

; Sequence 1796, Application US/10501282

; Publication No. US20050203280A1

; GENERAL INFORMATION:

; APPLICANT: MCMICHAEL, JOHN CALHOUN

; APPLICANT: ZAGURSKY, ROBERT JOHN

; APPLICANT: RUSSELL, DAVID PARRISH

; APPLICANT: FLETCHER, LEAH DIANE

; TITLE OF INVENTION: ALLOIOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING

; TITLE OF INVENTION: POLYPEPTIDE ANTIGENS, IMMUNOGENIC COMPOSITIONS AND USES THEREOF

; FILE REFERENCE: AM100780 L2

; CURRENT APPLICATION NUMBER: US/10/501,282

; CURRENT FILING DATE: 2004-07-09

; PRIOR APPLICATION NUMBER: 60/333,777

; PRIOR FILING DATE: 2001-11-29

; PRIOR APPLICATION NUMBER: 60/426,742

; PRIOR FILING DATE: 2002-11-18

; PRIOR APPLICATION NUMBER: PCT/US02/36123
; PRIOR FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1796
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Alloiococcus otitidis
US-10-501-282-1796

Query Match 32.6%; Score 153; DB 5; Length 334;
Best Local Similarity 39.1%; Pred. No. 1.6e-09;
Matches 36; Conservative 19; Mismatches 31; Indels 6; Gaps 4;

QY 1 LQGFVVPVYP-SSISELKGDKFKNK-MIGIDAGAGTQIVTQALNYYGLSKEYELV 58
Db 105 LGLVVPFYNEGDSIEDLE---DQIEDKTIIVGIERGAGTVEAERALEDYG-KEDWEVI 160
QY 59 PSSESVMLASLDSSIKRNEWILVPLWKPHWF 90
Db 161 ASSSGAMVTELQAINNEEDIIVTGTWTPHKF 192

RESULT 10
US-10-501-282-1798
; Sequence 1798, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOIOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; TITLE OF INVENTION: POLYPEPTIDE ANTIGENS, IMMUNOGENIC COMPOSITIONS AND USES THEREOF
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; PRIOR FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; PRIOR FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1798
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Alloiococcus otitidis
US-10-501-282-1798

Query Match 32.6%; Score 153; DB 5; Length 352;
Best Local Similarity 39.1%; Pred. No. 1.7e-09;
Matches 36; Conservative 19; Mismatches 31; Indels 6; Gaps 4;

QY 1 LQGFVVPVYP-SSISELKGDKFKNK-MIGIDAGAGTQIVTQALNYYGLSKEYELV 58
Db 123 LGLVVPFYNEGDSIEDLE---DQIEDKTIIVGIERGAGTVEAERALEDYG-KEDWEVI 178
QY 59 PSSESVMLASLDSSIKRNEWILVPLWKPHWF 90
Db 179 ASSSGAMVTELQAINNEEDIIVTGTWTPHKF 210

RESULT 11
US-10-501-282-1792
; Sequence 1792, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE

; TITLE OF INVENTION: ALLOIOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; TITLE OF INVENTION: POLYPEPTIDE ANTIGENS, IMMUNOGENIC COMPOSITIONS AND USES THEREOF
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; PRIOR FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; PRIOR FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1792
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Alloiococcus otitidis
US-10-501-282-1792

Query Match 32.4%; Score 152.5; DB 5; Length 320;
Best Local Similarity 36.4%; Pred. No. 1.7e-09;
Matches 32; Conservative 17; Mismatches 36; Indels 3; Gaps 1;

QY 3 GFVVPVYPISSISELKGDKFKNKMTIGIDAGAGTQIVTQALNYYGLSKEYELVPSSE 62
Db 122 GLGVPTYMEVDSIADLT--DEAEQTITIGIEAGAGVVGQALEDYDNLSDWTATSSS 178
QY 63 SVMLASLDSSIKRNEWILVPLWKPHWF 90
Db 179 GAMVTELETAINNEEDIIVTAWSPHKF 206

RESULT 12
US-09-815-242-5213
; Sequence 5213, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5213
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5213

Query Match 29.4%; Score 138; DB 3; Length 312;


```
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49743
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Burkholderia fungorum
US-10-282-122A-49743

Query Match      27.0%; Score 127; DB 4; Length 380;
Best Local Similarity 30.3%; Pred. No. 2.5e-06;
Matches 27; Conservative 24; Mismatches 32; Indels 6; Gaps 3;

QY      4 FVVPYSY---VPISSISELKGDKFKNKMIGIDAGAGTQIVTEQ--ALNYYGLSKEYELV 58
Db      184 FAVPSVEYDAGLKTADIAKHRDQLDGLGKIYGIPEGSSANAAIQKWTASNQFGL-RGFKLI 242

QY      59 PSSESVMLASLDSSIKRNEWILVPLWKPH 87
Db      243 ESSEAGMLVSDRAIREKKWVFLGWEPH 271
```

Search completed: January 24, 2006, 20:54:29
Job time : 16.9938 secs

; CURRENT APPLICATION NUMBER: US/10/942,072
 ; CURRENT FILING DATE: 2004-09-14
 ; PRIOR APPLICATION NUMBER: US/09/750,240
 ; PRIOR FILING DATE: 2001-10-12
 ; PRIOR APPLICATION NUMBER: US 09/472,667
 ; PRIOR FILING DATE: 1999-12-27
 ; PRIOR APPLICATION NUMBER: US 09/008,097
 ; PRIOR FILING DATE: 1998-01-16
 ; PRIOR APPLICATION NUMBER: US 08/924,757
 ; PRIOR FILING DATE: 1997-09-05
 ; PRIOR APPLICATION NUMBER: US 60/048,933
 ; PRIOR FILING DATE: 1997-06-16
 ; PRIOR APPLICATION NUMBER: US 08/708,661
 ; PRIOR FILING DATE: 1996-09-05
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13
 ; LENGTH: 1167
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Modified AC-VI
 US-10-942-072-13

Query Match 13.8%; Score 65; DB 6; Length 1167;
 Best Local Similarity 26.1%; Pred. No. 14;
 Matches 23; Conservative 10; Mismatches 25; Indels 30; Gaps 3;
 QY 32 IDAGAGTQI-VTEQALNYGLSKYELVP-----SSES 63
 Db 505 MEAGGRRITRATIQY--LNGDYVEPCRGGERNAVLYKEQIETFLILGASQKRKEK 562
 QY 64 VMLASLSSIKNEWILVPLKPHWAFS 91
 Db 563 AMLAKLQRTANSMEGLMPRWVPDRAFS 590

RESULT 3
 US-11-127-877-74
 ; Sequence 74, Application US/11/27877
 ; Publication No. US20050287565A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hoffmann, Marcel
 ; APPLICANT: Merchiers, Pascal G.
 ; APPLICANT: Spittaels, Koenraad F. F.
 ; APPLICANT: Laenen, Wendy
 ; TITLE OF INVENTION: Methods, Compositions and Compound Assays For Inhibiting
 ; FILE REFERENCE: P27,800-B USA
 ; CURRENT APPLICATION NUMBER: US/11/127,877
 ; PRIOR FILING DATE: 2005-05-12
 ; PRIOR APPLICATION NUMBER: 60/570,352
 ; PRIOR FILING DATE: 2004-05-12
 ; PRIOR APPLICATION NUMBER: 60/603,948
 ; PRIOR FILING DATE: 2004-08-24
 ; NUMBER OF SEQ ID NOS: 590
 ; SOFTWARE: Patent in version 3.3
 ; SEQ ID NO 74
 ; LENGTH: 837
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-11-127-877-74

Query Match 13.5%; Score 63.5; DB 7; Length 837;
 Best Local Similarity 31.9%; Pred. No. 14;
 Matches 23; Conservative 8; Mismatches 22; Indels 19; Gaps 3;
 QY 13 SSISLKGKGDYKFK---NKMIGIDACAGTQIVTEQALN-----YGLSKE 54
 Db 683 SCGSKGSGFRKPRYNNVVTIPAGA--THLVROQGNPHRSIYLALKLPDGSYALNGE 741
 QY 55 YELVPSSSEVML 66

Db 742 YTLMPSPDVL 753
 RESULT 4
 US-10-518-341-1
 ; Sequence 1, Application US/10518341
 ; Publication No. US20050249744A1
 ; GENERAL INFORMATION:
 ; APPLICANT: VAN ELS, Cecile Antoinette Carola Maria
 ; APPLICANT: VAN BLEEK, Gerarda Maria
 ; APPLICANT: BOOG, Marie-Claire Josefa Pia
 ; TITLE OF INVENTION: MHC class II haplotype specific immunodominancy of
 ; TITLE OF INVENTION: peptides derived from RSV fusion (F) or attachment
 ; TITLE OF INVENTION: (G) proteins
 ; FILE REFERENCE: 28902.0014
 ; CURRENT APPLICATION NUMBER: US/10/518,341
 ; CURRENT FILING DATE: 2004-12-17
 ; PRIOR APPLICATION NUMBER: PCT/NL03/00454
 ; PRIOR FILING DATE: 2003-06-20
 ; PRIOR APPLICATION NUMBER: EP 02077461.8
 ; PRIOR FILING DATE: 2002-06-20
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 574
 ; TYPE: PRT
 ; ORGANISM: respiratory syncytial virus
 US-10-518-341-1

Query Match 12.8%; Score 60; DB 6; Length 574;
 Best Local Similarity 20.0%; Pred. No. 21;
 Matches 16; Conservative 25; Mismatches 29; Indels 10; Gaps 3;
 QY 3 GFVPSYVPISSISSELKKGDKFKNMIG-----IDAGAGTQIVTEQAL---NYGLSKE 54
 Db 145 GSAIASGVAVSKVLHLEGEVKNKISALLSTNKAVVSLNSGVSLTSKVLDLKNY--IDKQ 202
 QY 55 YELVPSSSEVMLASLSSIK 74
 Db 203 LLPVINKQSCSISNIETVIE 222

RESULT 5
 US-11-022-562-214
 ; Sequence 214, Application US/11022562
 ; Publication No. US20050249742A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruprecht, Ruth M.
 ; APPLICANT: Shisong, Jiang
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING
 ; TITLE OF INVENTION: A CYTOTOXIC T LYMPHOCYTE IMMUNE RESPONSE
 ; FILE REFERENCE: DFN-043CN
 ; CURRENT APPLICATION NUMBER: US/11/022,562
 ; CURRENT FILING DATE: 2004-12-22
 ; PRIOR APPLICATION NUMBER: PCT/US03/20322
 ; PRIOR FILING DATE: 2003-06-27
 ; PRIOR APPLICATION NUMBER: 60/392718
 ; PRIOR FILING DATE: 2002-06-27
 ; NUMBER OF SEQ ID NOS: 340
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 214
 ; LENGTH: 574
 ; TYPE: PRT
 ; ORGANISM: Respiratory Syncytial Virus
 US-11-022-562-214

Query Match 12.8%; Score 60; DB 7; Length 574;
 Best Local Similarity 20.0%; Pred. No. 21;
 Matches 16; Conservative 25; Mismatches 29; Indels 10; Gaps 3;
 QY 3 GFVPSYVPISSISSELKKGDKFKNMIG-----IDAGAGTQIVTEQAL---NYGLSKE 54
 Db 145 GSAIASGVAVSKVLHLEGEVKNKISALLSTNKAVVSLNSGVSLTSKVLDLKNY--IDKQ 202


```

Db      539 GFYVATVV-----AKDRDVGDGVKVEHSLMGNAVSYLIRIDKETGEIFVTENEAFNYHRQNEL 593
QY      55 YELVPSSSEV 64
       : :|: :|:
Db      594 FVQIPADDTL 603

RESULT 8
US-11-193-715-4
; Sequence 4, Application US/11193715
; Publication No. US20050260675A1
; GENERAL INFORMATION:
; APPLICANT: Flannagan, Ronald D.
; APPLICANT: Mathis, John P.
; APPLICANT: Meyer, Terry E.
; TITLE OF INVENTION: Novel Bt Toxin Receptors From
; TITLE OF INVENTION: Lepidopteran Insects and Methods of Use
; FILE REFERENCE: 35718/204664
; CURRENT APPLICATION NUMBER: US/11/193,715
; CURRENT FILING DATE: 2005-07-29
; PRIOR APPLICATION NUMBER: US/09/715,909
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,285
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1730
; TYPE: PRT
; ORGANISM: Heliothis zea
US-11-193-715-4

Query Match      12.4%; Score 58.5; DB 7; Length 1730;
Best Local Similarity 25.7%; Pred. No. 1.3e+02;
Matches 18; Conservative 18; Mismatches 21; Indels 13; Gaps 3

QY      3 GFVPVSYPVSISELKKGDKFKNMIG-----IDAGAGTQIVTE-QALNYGLSLKE 54
       ||| :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db      539 GFYVATVV-----AKDRDVGDGVKVEHSLMGNAVSYLIRIDKETGEIFVTENEAFNYHRQNEL 593
QY      55 YELVPSSSEV 64
       : :|: :|:
Db      594 FVQIPADDTL 603

RESULT 9
US-11-194-246-303
; Sequence 303, Application US/11194246
; Publication No. US20050272089A1
; GENERAL INFORMATION:
; APPLICANT: Mott, John
; APPLICANT: Trepod, Catherine
; APPLICANT: Arvidson, Staffan
; TITLE OF INVENTION: CRITICAL GENES AND POLYPEPTIDES OF HAEMOPHILUS INFLUENZAE
; TITLE OF INVENTION: USE
; FILE REFERENCE: 00592.US1 (W&R 268.05920101)
; CURRENT APPLICATION NUMBER: US/11/194,246
; CURRENT FILING DATE: 2005-08-01
; PRIOR APPLICATION NUMBER: US/10/274,586
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US 60/345,438
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 621
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 303
; LENGTH: 318
; TYPE: PRT
; ORGANISM: HAEMOPHILUS INFLUENZAE
US-11-194-246-303

Query Match      12.3%; Score 58; DB 7; Length 318;
Best Local Similarity 19.8%; Pred. No. 1.7;
Matches 17; Conservative 20; Mismatches 33; Indels 16; Gaps 3

```

177 RRDSFRAEKILIDRLYKKVEEGKIVL 202

```

RESULT 12
US-11-165-226-53
; Sequence 53, Application US/11165226
; Publication No. US20050287627A1
; GENERAL INFORMATION:
; APPLICANT: KYOWA HAKKO KOGYO CO., LTD.
; TITLE OF INVENTION: Process for producing dipeptides or dipeptide derivatives
; FILE REFERENCE: 4093-13
; CURRENT APPLICATION NUMBER: US/11/165,226
; CURRENT FILING DATE: 2005-06-24
; PRIOR APPLICATION NUMBER: JP2004-189007
; PRIOR FILING DATE: 2004-06-25
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 1078
; TYPE: PRT
; ORGANISM: Brevibacillus brevis
US-11-165-226-53

```

```

Query Match      12.3%; Score 58; DB 7; Length 1078;
Best Local Similarity 24.7%; Pred. No. 82;
Matches 18; Conservative 19; Mismatches 26; Indels 10; Gaps 3;

Qy 18 LKGGDKDFKNMIGIDAGAGTQIVTEQA-----LNNY----GLSKEYELVPSSSE-SVMMLA 67
|||
Db 763 LPRTSDSPFDWSTLELFKYANSLFLFEAEAYWHHLNNTYENVOIKDYVTMNKKOINRYV 822

```

```

US-10-606-302-9
: Sequence 9, Application US/10606302
: Publication No. US20050260722A1
: GENERAL INFORMATION:
: APPLICANT: Porro, Danilo
: APPLICANT: Sauer, Michael
: TITLE OF INVENTION: Ascorbic Acid Production from Yeast

```

```

US-10-606-302-9
; Sequence 9, Application US/10606302
; Publication No. US20050260722A1
; GENERAL INFORMATION:
; APPLICANT: Porro, Danilo
; APPLICANT: Sauer, Michael
; TITLE OF INVENTION: Ascorbic Acid Production from Yeast
; FILE REFERENCE: 2028.594000
; CURRENT APPLICATION NUMBER: US/10/606,302
; CURRENT FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-606-302-9

```

```

Query Match      12.2%; Score 57.5; DB 6; Length 440;
Best Local Similarity 44.8%; Pred. No. 29;
Matches 13; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

Qy    59 PSSESVMLASLDSSIKRNEWILVPLWKPH 87
      ||| : : ||| : ||| : |||
Db    196 PSTLKEVLNDLSHLKARSEYFRF-LWFPH 223

RESULT 14
US-10-942-072-11
; Sequence 11, Application US/10942072
; Publication No. US20050250721A1
; GENERAL INFORMATION:
; APPLICANT: HAMMON, H. K.
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:28:54 ; Search time 3.45228 Seconds
(without alignments)
2536.214 Million cell updates/sec

Title: US-10-688-058-6
Perfect score: 470
Sequence: 1 LQGFVPSYVPISSELK.....SIKNEWILVPLWKPWFAPS 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	468	99.6	290	2 H70117	glycine betaine, L
2	193	41.1	285	2 H83241	probable glycine b
3	186.5	39.7	573	2 C86806	hypothetical prote
4	176	37.4	292	2 AI3042	hypothetical prote
5	176	37.4	304	2 C98243	hypothetical prote
6	175.5	37.3	287	2 AE3578	glycine betaine/l-
7	138	29.4	312	2 H82972	hypothetical prote
8	127	27.0	312	2 C82374	hypothetical prote
9	127	27.0	312	2 D95891	probable glycine-b
10	126	26.8	307	2 D83642	hypothetical prote
11	124.5	26.5	233	2 I40537	glycine betaine AB
12	124.5	26.5	300	2 AF1559	glycine betaine AB
13	124.5	26.5	300	2 AH1201	glycine betaine AB
14	108	23.0	308	2 AE2829	hypothetical prote
15	108	23.0	337	2 C97607	hypothetical prote
16	104	22.1	326	2 AC3307	glycine betaine/l-
17	101	21.5	317	2 AC3128	conserved hypotet
18	101	21.5	363	2 F98159	hypothetical prote
19	88	18.7	314	2 H83321	conserved hypotet
20	82.5	17.6	265	2 E97633	hypothetical prote
21	82.5	17.6	317	2 AH2856	hypothetical prote
22	80	17.0	331	2 AC0842	glycine betaine-bi
23	76.5	16.3	903	2 F82080	preproteins transi
24	76	16.2	322	2 E83009	probable binding p
25	75	16.2	327	2 H59094	hypothetical prote
26	74	15.7	319	2 F84966	thioredoxin-disulf
27	73.5	15.6	687	2 T04327	probable serine/th
28	73	15.5	334	2 AG0322	glycine betaine-bi
29	72	15.3	445	2 G70371	UDP-MURNAC-pentape

30	71.5	15.2	356	2 JX0141	cytochrome-c oxida
31	69.5	14.8	214	2 S46476	cysteine proteinas
32	69	14.7	467	2 S61141	probable membrane
33	68.5	14.6	345	2 AG3186	hypothetical prote
34	68	14.5	299	2 AI1208	metal binding prote
35	67.5	14.4	1182	2 E97286	DNA-dependent RNA
36	67.5	14.4	1199	2 G69698	RNA polymerase (be
37	67.5	14.4	1393	2 B86501	RNA polymerase bet
38	67.5	14.4	1393	2 E72122	DNA polymerase bet
39	67.5	14.4	1396	2 F81686	DNA-directed RNA p
40	67.5	14.4	1396	2 G71529	DNA-directed RNA p
41	67.5	14.4	1397	2 B81548	DNA-directed RNA p
42	67	14.3	562	2 AD0304	probable hemolysin
43	67	14.3	1005	2 T31333	beta-galactosidase
44	66	14.0	270	2 A82245	periplasmic bindin
45	66	14.0	323	2 B83215	conserved hypotet

ALIGNMENTS

RESULT 1

H70117
glycine betaine, L-proline ABC transporter, glycine/betaine/L-proline-binding protein (L
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C:Accession: H70117
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Whit-
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: H70117
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-290 <KLE>
A:Cross-references: UNIPROT:O51169; UNIPARC:UPI000005737B; GB:AE001125; GB:AE000783; NIT
A:Experimental source: strain B31

Query Match	99.6%	Score 468;	DB 2;	Length 290;
Best Local Similarity	98.9%	Pred. No. 1.1e-41;		
Matches	90;	Conservative	1;	Mismatches 0;
				Indels 0;
				Gaps 0;
QY	1	LQGFVPSYVPISSELKGGKDKFNKMIGIDAGAGTQIVTEQALNYGLSKVELYPS	60	
		;		
Db	109	IQGFVPSYVPISSELKGGKDKFNKMIGIDAGAGTQIVTEQALNYGLSKVELYPS	168	
QY	61	SESVMASLDSSIKRNEWILVPLWKPWFAPS	91	
		;		
Db	169	SESVMASLDSSIKRNEWILVPLWKPWFAPS	199	

RESULT 2

H83241
probable glycine betaine-binding protein precursor PA3236 [imported] - Pseudomonas aeru
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: H83241
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lam,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: H83241
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-285 <STO>
A:Cross-references: UNIPROT:Q9HZ04; UNIPARC:UPI00000C5968; GB:AE004746; GB:AE004091; NIT
A:Experimental source: strain PA01
C:Genetics:

A:Gene: PA3236

Query Match 41.1%; Score 193; DB 2; Length 285;
Best Local Similarity 42.7%; Pred. No. 9.7e-13;
Matches 38; Conservative 20; Mismatches 29; Indels 2; Gaps 1;

QY 3 GFVVPSPVPISSISELKGDKGKFNKMGIDAGAGTQIVTEQALNYGLSKEYELVPSS 62
DB 116 GLIVPEYVSANSIADLQAKDAFGGRVVGIDAGAGWIKTDEAIKQYGL--DYKLVAASG 173

QY 63 SVMLASLDSISIKRNEWILVPLWKPWFPS 91
DB 174 SGMIARELTAEKPKPVVTGWIPIHMF 202

RESULT 3

C86806
hypothetical protein busAB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: C86806
R:Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: C86806
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-573 <STO>
A:Cross-references: UNIPROT:O9KIP6; UNIPARC:UPI00000C611D; GB:AE005176; PID:gl2724442; F
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: busAB

Query Match 39.7%; Score 186.5; DB 2; Length 573;
Best Local Similarity 40.4%; Pred. No. 1.1e-11;
Matches 36; Conservative 21; Mismatches 29; Indels 3; Gaps 1;

QY 3 GFVVPSPVPISSISELKGDKGKFNKMGIDAGAGTQIVTEQALNYGLSKEYELVPSS 62
DB 406 GFVVPSPVSYMNVSIEDLTNOANK---TITGIEPGAGVMAASEKTLNSYDLNLDKWLVPSS 462

QY 63 SVMLASLDSISIKRNEWILVPLWKPWFPS 91
DB 463 GATVALGEAIKQKIDVITGWSPHWMFN 491

RESULT 4

A13042
hypothetical protein Atu3957 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: A13042
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: A13042
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-292 <KUR>
A:Cross-references: UNIPROT:Q8U8Y1; UNIPARC:UPI00001648AD; GB:AE008689; PIDN:AAL44759.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu3957
A:Map position: linear chromosome

Query Match 37.4%; Score 176; DB 2; Length 292;
Best Local Similarity 41.4%; Pred. No. 6.2e-11;

Matches 36; Conservative 14; Mismatches 35; Indels 2; Gaps 1;

QY 2 QGVVPSPVPISSISELKGDKGKFNKMGIDAGAGTQIVTEQALNYGLSKEYELVPSS 61
DB 119 QAIAPVKYVTIDSMDQLNDNADKFGKGIIGIEPGSLMRDAGAKAADYGL--KLQIVEGS 176

QY 62 ESMVLASLDSISIKRNEWILVPLWKPWFPS 88
DB 177 TAAMTAALKSAIDRKWEVAVTVWPEPSW 203

RESULT 5

C98243
hypothetical protein AGR_L1793 [imported] - Agrobacterium tumefaciens (strain C58, Core
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: C98243
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: C98243
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-304 <KUR>
A:Cross-references: UNIPROT:Q8U8Y1; UNIPARC:UPI00000D23BB; GB:AE007870; PIDN:AAK89469.1;
C:Genetics:
A:Gene: AGR_L1793
A:Map position: linear chromosome

Query Match 37.4%; Score 176; DB 2; Length 304;
Best Local Similarity 41.4%; Pred. No. 6.5e-11;
Matches 36; Conservative 14; Mismatches 35; Indels 2; Gaps 1;

QY 2 QGVVPSPVPISSISELKGDKGKFNKMGIDAGAGTQIVTEQALNYGLSKEYELVPSS 61
DB 131 QAIAPVKYVTIDSMDQLNDNADKFGKGIIGIEPGSLMRDAGAKAADYGL--KLQIVEGS 188

QY 62 ESMVLASLDSISIKRNEWILVPLWKPWFPS 88
DB 189 TAAMTAALKSAIDRKWEVAVTVWPEPSW 215

RESULT 6

AE3578
glycine betaine/l-proline-binding protein proX [imported] - Brucella melitensis (strain
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AE3578
R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
; Mazur, M.; Goltzman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella meliten
A:Reference number: AD3252; PMID:11756688
A:Accession: AE3578
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-287 <KUR>
A:Cross-references: UNIPROT:Q8YCH8; UNIPARC:UPI0000058515; GB:AE008918; PIDN:AAL53792.1;
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME110550
A:Map position: II

Query Match 37.3%; Score 175.5; DB 2; Length 287;
Best Local Similarity 38.7%; Pred. No. 6.9e-11;
Matches 36; Conservative 21; Mismatches 29; Indels 7; Gaps 3;

QY 3 GFVVPSPVPISSISELKGK---ISSISELKGK--GDKPKNKMIGIDAGAGTQIVTEQALNYGLSKEYEL 57
DB 114 GWIVPDYTPESVSSIEDLKPEVREKLKGEIQGIDPGAGLTRLQAEAIKKYGL--DYKL 171

```
QY 58 VPSSSVMLASLDSSIKRNEWILVPLWKPHAF 90
Db 172 NISSEAMLTVDRAIRSEGFVATSWSPHMF 204

RESULT 7
H82972
hypothetical protein PA5388 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: H82972
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: H82972
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-312 <STO>
A:Cross-references: UNIPROT:Q9HTH6; UNIPARC:UPI00000C5FD9; GB:AE004951; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA5388

Query Match 29.4%; Score 138; DB 2; Length 312;
Best Local Similarity 33.7%; Pred. No. 6.8e-07;
Matches 30; Conservative 24; Mismatches 29; Indels 6; Gaps 3;

QY 4 FVPSYVYP---ISSISELKGKDKFKNMIGIDAGAGTQIVTEQAL--NYGLSKEYELV 58
Db 117 YAVPSYVAEGLGKTFADARFKDKLGGRIVYALEPGSGSNRTRKMDNRRFEL-KGFQLV 175

QY 59 PSSESVMLASLDSSIKRNEWILVPLWKPH 87
Db 176 ESSEAGMLTAVKRAIKRQWVFFGCKPH 204

RESULT 8
C82974
hypothetical protein PA5378 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: C82974
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: C82974
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-312 <STO>
A:Cross-references: UNIPROT:Q9HTI6; UNIPARC:UPI00000C5FCF; GB:AE004950; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA5378

Query Match 27.0%; Score 127; DB 2; Length 312;
Best Local Similarity 34.5%; Pred. No. 9.8e-06;
Matches 30; Conservative 20; Mismatches 31; Indels 6; Gaps 3;

QY 6 VPSYV---PISSISELKGKDKFKNMIGIDAGAGTQIVTEQAL--NYGLSKEYELVPS 60
Db 117 VPQYVDGLRSFADIAKPSDKLGNKIYGIIEPGNGNRVAQSMIDKNFELGK-PKLVES 175

QY 61 SESVVMLASLDSSIKRNEWILVPLWKPH 87
Db 176 SEAGMLSQVQAIRNQWVFFLGNEPH 202
```

```
RESULT 9
D95891
probable glycine-betaine-binding protein [imported] - Sinorhizobium meliloti (strain 102
C:Species: Sinorhizobium meliloti
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: D95891
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A>Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: D95891
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-312 <KUR>
A:Cross-references: UNIPROT:Q92WE4; UNIPARC:UPI00000CB4FA; GB:AL591985; PIDN:CAC48796.1.;
A:Experimental source: strain 1021, megaplasmid pSymB
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebauchot, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wellis, D.H.; Wong, K.; Yeh, K.
A>Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMB20410
A:Genome: plasmid

Query Match 27.0%; Score 127; DB 2; Length 312;
Best Local Similarity 37.1%; Pred. No. 9.8e-06;
Matches 33; Conservative 16; Mismatches 33; Indels 6; Gaps 3;

QY 4 FVPSYVYP---ISSISELKGKDKFKNMIGIDAGAGTQIVTEQALNY--YGLSKEYELV 58
Db 115 FAVPKYVADAGVKDISDLQKFPDKFRKIYGIIEPGNGNRMLDMINKGDFGLT-GWELV 173

QY 59 PSSESVMLASLDSSIKRNEWILVPLWKPH 87
Db 174 ESSEQGLAERATKDDQWIVFLGWAPH 202

RESULT 10
D83642
hypothetical protein PA0030 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: D83642
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: D83642
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-307 <STO>
A:Cross-references: UNIPROT:Q9I7A0; UNIPARC:UPI00000C4EF4; GB:AE004442; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0030

Query Match 26.8%; Score 126; DB 2; Length 307;
Best Local Similarity 34.5%; Pred. No. 1.2e-05;
Matches 30; Conservative 16; Mismatches 35; Indels 6; Gaps 3;

QY 6 VPSYV---PISSISELKGKDKFKNMIGIDAGAGTQIVTEQAL--NYGLSKEYELVPS 60
Db 117 VPRIYVDAGVHRFEDLAAQGRFNRKLYIGISGAPANQSIQKMDANQFGLG-DWKLVES 175

QY 61 SESVVMLASLDSSIKRNEWILVPLWKPH 87
```

Db 176 SEQAMLAELGRAEKQRMLVFLGWTFH 202

RESULT 11

glycine betaine ABC transporter (glycine betaine-binding protein) opuAC precursor - Bacillus subtilis

C:Species: *Bacillus subtilis*

C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004

C:Accession: I40537; F59669

R:Kempf, B.; Bremer, E.

J. Biol. Chem. 270, 16701-16713, 1995

A:Title: OpuA, an osmotically regulated protein-dependent transport system for betaine

A:Reference number: A57322; PMID:95348093; PMID:7622480

A:Accession: I40537

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-293 <RES>

A:Cross-references: UNIPROT:P46922; UNIPARC:UPI000005FED3; EMBL:U17292; NID:9984802; PID:140537

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteaux, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chao, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maugda, S.; Maueel, Y.; M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanton, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

A:Reference number: A69580; PMID:98044033; PMID:9384377

A:Accession: F63669

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-293 <KUN>

A:Cross-references: UNIPARC:UPI000005FED3; GB:Z99105; GB:AL009126; NID:92632457; PIDN:CA

A:Experimental source: strain 168

C:Genetics:

A:Gene: opuAC

Query Match 26.5%; Score 124.5; DB 2; Length 293;

Best Local Similarity 32.9%; Pred. No. 1.7e-05;

Matches 26; Conservative 18; Mismatches 34; Indels 1; Gaps 1;

QY 13 SSTSELKRGDKPKFKNMIGIDAGAGTQIVTEQALNYGYLSKEVELVPSSSEVMLASLDSS 72

Db 23 SENDENASAAEQVNKTIIGIDPGSGIMSLTDKMKDYDLN-DWTLLSASSAAWTATLKS 81

QY 73 IKRNEWILVPLWKPHWAPS 91

Db 82 YDRKKPIITGTWTFHWMS 100

RESULT 12

AF1559

glycine betaine ABC transporters (glycine betaine-binding protein) homolog gbuC [imported]

C:Species: *Agrobacterium tumefaciens*

C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C:Accession: AF1559

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria species*.

A:Reference number: AB1077; PMID:21537279; PMID:11679669

A:Accession: AF1559

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-308 <KUR>

A:Residues: 1-300 <GLA>

A:Cross-references: UNIPROT:Q92D06; UNIPARC:UPI000000CC423; GB:AL592022; PIDN:CAC96246.1;

A:Experimental source: strain Clp11262

C:Genetics:

A:Gene: gbuC

Query Match 26.5%; Score 124.5; DB 2; Length 300;

Best Local Similarity 35.1%; Pred. No. 1.7e-05;

Matches 26; Conservative 16; Mismatches 31; Indels 1; Gaps 1;

QY 19 KGKGDKFKNMIGIDAGAGTQIVTEQALNYGYLSKE-YELVPSSSEVMLASLDSSIKRNE 77

Db 32 KDLGEQINYYTITGDAGAGIMLATQNAIKDYHLDDNNWLOQTSSSTAAMTSTLQKAMKDKR 91

QY 78 WILVPLWKPHWAPS 91

Db 92 PIVVTGWTWTFHWMT 105

RESULT 13

AH1201

glycine betaine ABC transporters (glycine betaine-binding protein) homolog gbuC [imported]

C:Species: *Listeria monocytogenes*

C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C:Accession: AH1201

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria species*.

A:Reference number: AB1077; PMID:21537279; PMID:11679669

A:Accession: AH1201

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-300 <GLA>

A:Cross-references: UNIPROT:Q8Y898; UNIPARC:UPI000000CF12A; GB:NC_003210; PIDN:CAC99094.1;

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: gbuC

Query Match 26.5%; Score 124.5; DB 2; Length 300;

Best Local Similarity 35.1%; Pred. No. 1.7e-05;

Matches 26; Conservative 16; Mismatches 31; Indels 1; Gaps 1;

QY 19 KGKGDKFKNMIGIDAGAGTQIVTEQALNYGYLSKE-YELVPSSSEVMLASLDSSIKRNE 77

Db 32 KDLGEQINYYTITGDAGAGIMLATQNAIKDYHLDDNNWLOQTSSSTAAMTSTLQKAMKDKR 91

QY 78 WILVPLWKPHWAPS 91

Db 92 PIVVTGWTWTFHWMT 105

RESULT 14

AZ2829

hypothetical protein Atu2060 [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)

C:Species: *Agrobacterium tumefaciens*

C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

C:Accession: AZ2829

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, J.; erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A:Reference number: AB2577; PMID:21608550; PMID:11743193

A:Accession: AZ2829

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-308 <KUR>

Search completed: January 24, 2006, 19:59:01
Job time : 4.45228 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:27:53 ; Search time 19.5539 Seconds
(without alignments)
3283.383 Million cell updates/sec

Title: US-10-688-058-6

Perfect score: 470

Sequence: 1 LQGFVVPYSYPISSISLKG.....SIKNEWILVPLMKPHWAFS 91

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	468	99.6	290	051169 BORBU	051169 borrelia bu
2	459	97.7	294	066219 BORBU	066219 borrelia ga
3	218	46.4	277	08A6X6_BACTN	08A6X6 bacteroides
4	218	46.4	315	08TNY0_METAC	08TNY0 methanosarc
5	216.5	46.1	285	0729Q3_DESVH	0729Q3 desulfovibr
6	211.5	45.0	284	064R03_BACFR	064R03 bacteroides
7	210.5	44.8	284	05LAK8_BACFN	05LAK8 bacteroides
8	199	42.3	871	09RJ81_STRCO	09RJ81 streptomyce
9	194	41.3	287	063M02_BURPS	063M02 burkholderi
10	194	41.3	287	062DH6_BURMA	062DH6 burkholderi
11	193	41.1	285	09HZ04_PSEAE	09HZ04 pseudomonas
12	193	41.1	287	04LV54_9BURK	04LV54 burkholderi
13	192	40.9	871	0828F3_STRAM	0828F3 streptomyce
14	187.5	39.9	283	04KFZ7_PSEF5	04KFZ7 pseudomonas
15	186.5	39.7	573	09KIF6_9LACT	09KIF6 lactococcus
16	186.5	39.7	573	07DAU8_LACLA	07DAU8 lactococcus
17	183.5	39.0	306	08U4S3_METMA	08U4S3 methanosarc
18	180.5	38.4	573	09RQ05_9LACT	09RQ05 lactococcus
19	178	37.9	287	08M34_PSEPK	08M34 pseudomonas
20	177.5	37.8	285	05WCG0_BACSK	05WCG0 bacillus cl
21	176.5	37.6	288	08FTV7_BRUSU	08FTV7 brucella su
22	176	37.4	292	08U8Y1_AGRU5	08U8Y1 agrobacteri
23	176	37.4	304	07CTQ1_AGRU5	07CTQ1 agrobacteri
24	175.5	37.3	287	08YCH8_BRUME	08YCH8 brucella me
25	175.5	37.3	288	0578M1_BRUAB	0578M1 brucella ab
26	174	37.0	302	07V819_PROVM	07V819 prochloroco
27	172.5	36.7	285	063AG1_BACCZ	063AG1 bacillus ce
28	169	36.0	321	07U4Z5_SYNPX	07U4Z5 synechococ
29	169	36.0	328	0886F7_PSESM	0886F7 pseudomonas
30	168.5	35.9	285	06HHV0_BACHK	06HHV0 bacillus th
31	168.5	35.9	285	081PL5_BACAN	081PL5 bacillus an

32 168 35.7 289 2 Q8XTP2_RALSO Q8XTP2 ralstonia s
33 166 35.3 284 2 Q4ZPY3_PSESY Q4ZPY3 pseudomonas
34 166 35.3 295 2 Q6LYW9_METWP Q6LYW9 methanococc
35 166 35.3 298 2 Q891M9_CLOTE Q891M9 clostridium
36 165.5 35.2 285 2 Q4MKI4_BACCE Q4MKI4 bacillus ce
37 164 34.9 282 2 Q5WJD4_BACSK Q5WJD4 bacillus cl
38 163.5 34.8 285 2 Q81CH2_BACCR Q81CH2 bacillus ce
39 158.5 33.7 569 2 Q830X6_ENTFA Q830X6 enterococcu
40 157 33.4 575 2 Q9A1M9_STRPY Q9A1M9 streptococc
41 156 33.2 561 2 Q8K8R4_STRP3 Q8K8R4 streptococc
42 156 33.2 575 2 Q879N8_STRP3 Q879N8 streptococc
43 156 33.2 575 2 Q5XE24_STRP6 Q5XE24 streptococc
44 156 33.2 575 2 Q8P2S5_STRP8 Q8P2S5 streptococc
45 153 32.6 308 2 Q98HR5_RHILO Q98HR5 rhizobium 1

ALIGNMENTS

RESULT 1
051169 BORBU PRELIMINARY; PRT; 290 AA.
AC 051169
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Glycine betaine, L-proline ABC transporter, glycine/betaine/L-proline-
binding protein (ProX)
GN OrderedLocusNames=BB0144;
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=139;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Frazer C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Weidman J.F., Utterback T.R., Matthey L., McDonald L.A.,
RA Artach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi."
RL Nature 390:580-586(1997).
DR EMBL; AB001125; AAC6525.1; -; Genomic_DNA.
DR PIR; H70117; H70117.
DR TIGR; BB0144; -;
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007210; OpuAC_ABC.
DR Pfam; PF04069; OpuAC; 1.
KW Complete proteome.
SQ SEQUENCE 290 AA; 33263 MW; 43779CF0911FBBD5 CRC64;

Query Match 99.6%; Score 468; DB 2; Length 290;
Best Local Similarity 98.9%; Pred. No. 2,3e-40;
Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQGFVVPYSYPISSISLKGDKFKNMIGIDAGAGTQIVTEQALNYGLSKVELVPS 60
:|||||
Db 109 IQGFVVPYSYPISSISLKGDKFKNMIGIDAGAGTQIVTEQALNYGLSKVELVPS 168
:|||||
QY 61 SSSVNLASLDSSIKRNEWILVPLMKPHWAFS 91
:|||||
Db 169 SSSVNLASLDSSIKRNEWILVPLMKPHWAFS 199
:|||||

RESULT 2

Q662L9_BORGA

ID Q662L9_BORGA PRELIMINARY; PRT; 294 AA.

AC Q662L9;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 23-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Glycine betaine, L-proline ABC transporter, binding protein.

GN Name:proX; OrderedLocusNames=BG0144;

OS Borrelia garinii.

OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;

OC Borrelia burgdorferi group.

OX NCBI_TaxID=29519;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=PBI;

RA Gloeckner G., Lehmann R., Romualdi A., Pradella S.,

RT "Comparative analysis of the Borrelia garinii genome.";

RL Submitted (JUN-2004) to the EMBL/GenBank/DBJP databases.

DR EMBL; CP000013; AAU07002.1; -; Genomic_DNA.

DR GO; GO:0005488; F:binding; IEA.

DR GO; GO:0005215; F:transporter activity; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR007210; OpuAC_ABC.

DR Pfam; PF04069; OpuAC; 1.

KW Complete proteome.

SQ SEQUENCE 294 AA; 33680 MW; AFA81026194ABE12 CRC64;

Query Match 97.7%; Score 459; DB 2; Length 294;

Best Local Similarity 95.6%; Pred. No. 2e-39;

Matches 87; Conservative 4; Mismatches 0; Indels 0; Gaps 0

QY 1 LQGFVPSYVPISSISELKGKDFKNKMGIDGAGTQIVTEQALNYGLSKVELYVPS 60

DB 113 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 172

QY 61 SESVMLASLDSISIKRNEWILVPLMKPHWAFS 91

DB 173 SESVMLASLDSAIKRNWVILVPLMKPHWAFS 203

RESULT 3

ID Q8A6X6_BACTN PRELIMINARY; PRT; 277 AA.

AC Q8A6X6;

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Glycine betaine-binding protein.

GN OrderedLocusNames=Btl1749;

OS Bacteroides thetaiotaomicron.

OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;

OC Bacteroidaceae; Bacteroides.

OX NCBI_TaxID=818;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=VT-5482 / ATCC 29148;

EX MEDLINE=22550858; PubMed=12663928; DOI=10.1126/science.1080029;

RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,

RA Chiang H.C., Hooper L.V., Gordon J.I.;

RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";

RL Science 299:2074-2076(2003).

DR EMBL; AE016933; AA076856.1; -; Genomic_DNA.

DR GO; GO:0005488; F:binding; IEA.

DR GO; GO:0005215; P:transporter activity; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR007210; OpuAC_ABC.

DR Pfam; PF04069; OpuAC; 1.

KW Complete proteome.

SQ SEQUENCE 277 AA; 31056 MW; B020C694CE12A21A CRC64;

Query Match 46.4%; Score 218; DB 2; Length 277;

Best Local Similarity 47.7%; Pred. No. 2.7e-14;

Query Match	45.0%;	Score	211.5;	DB 2;	Length	284;			
Best Local Similarity	46.6%;	Pred. NO.	1.7e-13;						
Matches	41;	Conservative	16;	Mismatches	30;	Indels	1;	Gaps	1;
QY	3	GFVVPSPVPISSISLKGKDKFKNMIGIDAGAGTQIVTEQALNYGLSKSEYELVPSS	62						
Db	112	GLVVPQVTTIDSIGELAAHKQPSSEIVGIDAGAGIMKTTDKAIGDYGLD-GYKLLTSS	170						
QY	63	SVMLASLDSSIKRNEWILVPLWPKHPWAF	90						
Db	171	STWLASLQKAMEKEAWIVITGWTPHWPF	198						
<p>RESULT 7</p> <p>QSLAK8 BACFN</p> <p>ID QSLAK8 BACFN PRELIMINARY; PRT; 284 AA.</p> <p>AC QSLAK8;</p> <p>DT 01-FEB-2005 (TrEMBLrel. 29, Created)</p> <p>DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)</p> <p>DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)</p> <p>DE Putative exported glycine betaine-binding protein.</p> <p>GN OrderedLocusNames=BF3170;</p> <p>OS Bacteroides fragilis (strain ATCC 25285 / NCTC 9343).</p> <p>OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;</p> <p>OC Bacteroidaceae; Bacteroides.</p> <p>NCBI_TaxID=272559;</p> <p>[1]</p> <p>RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].</p> <p>RP PubMed=15746427; DOI=10.1126/science.1107008;</p> <p>RA Cardeno-Tarraga A.-M., Patrick S., Crossman L. C., Blakely G., Abratt V., Lennard N., Poxton I., Duerden B., Harris B., Quail M. A., Barron A., Clark L., Corton C., Doggett J., Holden M. T. G., Larke N., Line A., Lord A., Norbertczak H., Ormond D., Price C., Rabinowitz E., Woodward J., Barrell B. G., Parkhill J.</p> <p>RT "Extensive DNA inversions in the B. fragilis genome control variable gene expression."</p> <p>RL Science 307:1463-1465 (2005).</p> <p>DR EMBL; CR626927; CAH08865.1; -; Genomic_DNA.</p> <p>DR InterPro; IPR007210; OpuAC_ABC_sub_bd.</p> <p>DR Pfam; PF04069; OpuAC; 1.</p> <p>KW Complete proteome.</p> <p>SQ SEQUENCE 284 AA; 31962 MW; 4F34F719D44F0448 CRC64;</p>									
Query Match	44.8%;	Score	210.5;	DB 2;	Length	284;			
Best Local Similarity	46.6%;	Pred. NO.	1.7e-13;						
Matches	41;	Conservative	16;	Mismatches	30;	Indels	1;	Gaps	1;
QY	3	GFVVPSPVPISSISLKGKDKFKNMIGIDAGAGTQIVTEQALNYGLSKSEYELVPSS	62						
Db	112	GLVVPQVTTIDSIGELAAHKQPSSEIVGIDAGAGIMKTTDKAIGDYGLD-GYKLLTSS	170						
QY	63	SVMLASLDSSIKRNEWILVPLWPKHPWAF	90						
Db	171	STWLASLQKAMEKEAWIVITGWTPHWPF	198						
<p>RESULT 8</p> <p>Q9RJ81 STRCO</p> <p>ID Q9RJ81 STRCO PRELIMINARY; PRT; 871 AA.</p> <p>AC Q9RJ81;</p> <p>DT 01-MAY-2000 (TrEMBLrel. 13, Created)</p> <p>DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)</p> <p>DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)</p> <p>DE Glycine betaine transport system permease protein.</p> <p>GN OrderedLocusNames=SCO1620; ORFNames=SCI41.03c;</p> <p>OS Streptomyces coelicolor.</p> <p>OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;</p> <p>OC Streptomycineae; Streptomycetaceae; Streptomyces.</p> <p>NCBI_TaxID=1902;</p> <p>[1]</p> <p>RN NUCLEOTIDE SEQUENCE.</p> <p>RP STRAIN=A3(2) / M145;</p>									

```

RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hdaigo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RA "complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147 (2002).
CC -!- FUNCTION: Part of a binding-protein-dependent transport system.
CC Probably responsible for the translocation of the substrate across
CC the membrane (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the binding-protein-dependent transport
CC system permease family.
DR EMBL; AL939109; CAB59473.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD transp.
DR InterPro; IPR007210; OpuAC_ABC.
DR Pfam; PF00528; BPD_transp_I; 1.
DR Pfam; PF04069; OpuAC; 2.
DR PROSITE; PS05928; ABC_TM1; 1.
KW Complete proteome; Transmembrane; Transport.
SQ SEQUENCE 871 AA; 95028 MW; 4CC12B484AA43506 CRC64;

Query Match 42.3%; Score 199; DB 2; Length 871;
Best Local Similarity 46.6%; Pred. No. 9.3e-12;
Matches 41; Conservative 14; Mismatches 31; Indels 2; Gaps 2;

QY 6 VPSYVP-PISSISSELKGGKDFKKNMIGIDAGAG-TQIVTEQALNYVGLSKYELVPSSES 63
DB 699 VFSYVKDVKSLADLKGKDLFGKIGIGFGEVQQLKNVDLPVGYLEDEYVVDGSGTP 758

QY 64 VMLASLDSSIKRNEWILVPLWKPHWAPS 91
DB 759 AMLAEKRALAKKEPVAVTLMSPHWAYS 786

RESULT 9
Q63MU2 BURPS
ID Q63MU2 BURPS PRELIMINARY; PRT; 287 AA.
AC Q63MU2.
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Putative exported protein.
GN OrderedLocNames=BPSS0556;
OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; pseudomallei group.
OX NCBI_TaxId=28450;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K96243;
RX PubMed=15377794; DOI=10.1073/pnas.0403302101;
RA Holden M.T.G., Tibball R.W., Peacock S.J., Cerdano-Tarraga A.-M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.L.,
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crockett B., Davis P., Deshaizer D.,
RA Felwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitch E., Rutherford K., Sanders M., Simmonds M.,
RA Songvilailai S., Stevens K., Tunapa S., Vesaratchavest M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;


```

```

RT "Genomic plasticity of the causative agent of melioidosis,
RA Burkholderia pseudomallei.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245 (2004).
DR EMBL; BX571966; CAH38013.1; -; Genomic_DNA.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000583; GATase 2.
DR InterPro; IPR007210; OpuAC_ABC.
DR Pfam; PF04069; OpuAC; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 287 AA; 30851 MW; EF3DFA0B95DC82AB CRC64;

Query Match 41.3%; Score 194; DB 2; Length 287;
Best Local Similarity 42.7%; Pred. No. 9e-12;
Matches 38; Conservative 18; Mismatches 31; Indels 2; Gaps 1;

QY 3 GFVPSYVP-PISSISSELKGGKDFKKNMIGIDAGAGTQIVTEQALNYVGLSKYELVPSSE 62
DB 116 GLIVPAYVKAKTIADLNQAQKNDFAGRIVGIDAGAGWMKRTDDAIKAYGLN--YSLMPSSG 173

QY 63 SVMLASLDSSIKRNEWILVPLWKPHWAPS 91
DB 174 SAMTAELARSIHNSKAVVVTGWAPHWMPA 202

RESULT 10
Q62DH6 BURMA
ID Q62DH6 BURMA PRELIMINARY; PRT; 287 AA.
AC Q62DH6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE ABC transporter, periplasmic glycine/betaine-binding protein.
GN OrderedLocNames=BMAA0480;
OS Burkholderia mallei (Pseudomonas mallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxId=13373;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 23344;
RX PubMed=15377793; DOI=10.1073/pnas.0403306101;
RA Nieman W.C., Deshaizer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T.V., Ulrich R.L., Ronning C.M., Brinkac L.M.,
RA Daugherty S.C., Davidson T.D., DeBoy R.T., Dimitrov G., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Haft D.H., Khouri H.M., Kolonay J.F.,
RA Madupu R., Mohammed Y., Nelson W.C., Radune D., Romero C.M.,
RA Sarria S., Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y.,
RA Zafar N., Zhou L., Fraser C.M.;
RA "Structural flexibility in the Burkholderia mallei genome.";
RT Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251 (2004).
RL EMBL; CP000011; AAU46649.1; -; Genomic_DNA.
DR TIGR; BMAA0480; -;
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000583; GATase 2.
DR InterPro; IPR007210; OpuAC_ABC.
DR Pfam; PF04069; OpuAC; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 287 AA; 30851 MW; EF3DFA0B95DC82AB CRC64;

Query Match 41.3%; Score 194; DB 2; Length 287;
Best Local Similarity 42.7%; Pred. No. 9e-12;
Matches 38; Conservative 18; Mismatches 31; Indels 2; Gaps 1;

QY 3 GFVPSYVP-PISSISSELKGGKDFKKNMIGIDAGAGTQIVTEQALNYVGLSKYELVPSSE 62
DB 116 GLIVPAYVKAKTIADLNQAQKNDFAGRIVGIDAGAGWMKRTDDAIKAYGLN--YSLMPSSG 173


```



```
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD transp.
DR InterPro; IPR007210; OpuAC_ABC.
DR Pfam; PF00528; BPD transp_1; 1.
DR Pfam; PF04069; OpuAC; 2.
DR PROSITE; PS09328; ABC_TMI; 1.
KW Complete proteome; Transmembrane; Transport.
SQ SEQUENCE 871 AA; 94924 MW; E6004F3253C8925C CRC64;

Query Match 40.9%; Score 192; DB 2; Length 871;
Best Local Similarity 45.5%; Pred. No. 5e-11;
Matches 40; Conservative 15; Mismatches 31; Indels 2; Gaps 2;

QY 6 VPSYV-PISSISELKGDKFKNMIGIDAGAG-TQIVTEQALNYGLSKYELVPSSE 63
Db VPSYVVKVSLADLKGKGTDFGKIIGIEPTGEMNLLKTKVLPGLDKEYKVVGDGTP 758
QY 64 VMLASLDSISKRNEMILVPLWKPHWAFS 91
Db 759 AMLAEKRAYAKKEPVAVVLWSPHWAYS 786

RESULT 14
QAKFZ7_PSEFS
ID QAKFZ7_PSEFS PRELIMINARY; PRT; 283 AA.
AC QAKFZ7
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Glycine betaine/L-proline ABC transporter, glycine betaine/L-proline-
DE binding/permease protein.
GN Name=prox; ORFNames=PFL_1710;
OS Pseudomonas fluorescens (strain Pf-5).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=220664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Pf-5;
RX PubMed=15980861; DOI=10.1038/nbt1110;
RA Paulsen I.T., Press C., Ravel J., Kobayashi D., Myers G.S.,
RA Mavrodi D., DeBoy R.T., Seshadri R., Ren Q., Madupu R., Dodson R.J.,
RA Durkin S., Brinkac L.M., Daugherty S.C., Sullivan S.A., Rosovitz M.,
RA Gwinn M.L., Zhou L., Nelson W.C., Weidman J., Watkins K., Tran K.,
RA Khouri H.M., Pierson E., Pierson L. III, Thomasow L., Loper J.;
RT "Complete genome sequence of the plant commensal Pseudomonas
RT fluorescens Pf-5."
RL Nat. Biotechnol. 23:873-878(2005).
DR EMBL; CP000076; AAY91005.1; -; Genomic DNA.
SQ SEQUENCE 283 AA; 30765 MW; 85963158ED83E482 CRC64;

Query Match 39.9%; Score 187.5; DB 2; Length 283;
Best Local Similarity 43.8%; Pred. No. 4.2e-11;
Matches 39; Conservative 17; Mismatches 30; Indels 3; Gaps 2;

QY 3 GFVVPSPYVPISSISELKGDKFKNMIGIDAGAGTQIVTEQALNYGLSKYELVPSSE 62
Db 114 GLIVPEYVKAKSIDLK-TDTTFKNKIVGDAGSGVWLKTDQAIKDYGL--DYKLQASSG 170
QY 63 SVMLASLDSISKRNEMILVPLWKPHWAFS 91
Db 171 AAMIAELTRAEKQDSIAVTGWVPHMFA 199

RESULT 15
Q9KIF6_9LACT
ID Q9KIF6_9LACT PRELIMINARY; PRT; 573 AA.
AC Q9KIF6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
```

```
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE OpuABC.
CN Name=opuABC;
OS Lactococcus lactis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1358;
RN [1]
RP NUCLEOTIDE SEQUENCE.
EX MEDLINE=20318987; PubMed=10860977; DOI=10.1073/pnas.97.13.7102;
RA van der Heide T., Poolman B.;
RT "Osmoregulated ABC-transport system of Lactococcus lactis senses water
RT stress via changes in the physical state of the membrane.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:7102-7106(2000).
DR EMBL; AF234619; AAF37879.1; -; Genomic_DNA.
DR PIR; C86806; C86806.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0015450; F:protein translocase activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0009306; P:protein secretion; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD transp.
DR InterPro; IPR007210; OpuAC_ABC.
DR InterPro; IPR002208; SecY.
DR Pfam; PF00528; BPD transp_1; 1.
DR Pfam; PF04069; OpuAC; 1.
DR PRINTS; PR00303; SECYTRNLCASE.
DR PROSITE; PS0928; ABC_TMI; 1.
SQ SEQUENCE 573 AA; 61952 MW; DFB59CEAB8462A0F CRC64;

Query Match 39.7%; Score 186.5; DB 2; Length 573;
Best Local Similarity 40.4%; Pred. No. 1.2e-10;
Matches 36; Conservative 21; Mismatches 29; Indels 3; Gaps 1;

QY 3 GFVVPSPYVPISSISELKGDKFKNMIGIDAGAGTQIVTEQALNYGLSKYELVPSSE 62
Db 406 GFVVPSPYVNVNSIEDLTNQANK---TITGIEPGAGVMAASEKTLNSYDNLKDWKLVPSSS 462
QY 63 SVMLASLDSISKRNEMILVPLWKPHWAFS 91
Db 463 GAMTVLGEALIKQHKDVIITGNSPHWMFN 491

Search completed: January 24, 2006, 19:56:43
Job time : 22.5539 secs
```


SQ Sequence 290 AA;
 Query Match 100.0%; Score 1527; DB 2; Length 290;
 Best Local Similarity 100.0%; Pred. No. 5e-137;
 Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 MYKFLFFIIFMFLSCDEKSKSNLKVIGYVNWGGETAATNVLKVFEKMGYNAEIFS 60
 1 MYKFLFFIIFMFLSCDEKSKSNLKVIGYVNWGGETAATNVLKVFEKMGYNAEIFS 60
 61 VTTISIMYOYLAGSKIDGTSSWVPTADKFFYEKLTKEFVDLGNAYEGTIQGFVPSYVPI 120
 61 VTTISIMYOYLAGSKIDGTSSWVPTADKFFYEKLTKEFVDLGNAYEGTIQGFVPSYVPI 120
 121 SSISLKGKGDGFKNKMGIDAGAGTQIVTEQALNYGLSKEYELVPSSSVMLASLDSS 180
 121 SSISLKGKGDGFKNKMGIDAGAGTQIVTEQALNYGLSKEYELVPSSSVMLASLDSS 180
 181 IKRNEWILVPLWKPHWAFSRDYIKFLDDPDLIMGGIESVHTLVRLGLENDDFDAYVFDH 240
 181 IKRNEWILVPLWKPHWAFSRDYIKFLDDPDLIMGGIESVHTLVRLGLENDDFDAYVFDH 240
 241 FYWSDDLILPLMDKNDKEPKGEYRNAVEFVEKKEIVKTWPEKYKTLFD 290
 241 FYWSDDLILPLMDKNDKEPKGEYRNAVEFVEKKEIVKTWPEKYKTLFD 290
 RESULT 2
 ADU98749
 ID ADU98749 standard; protein; 290 AA.
 AC ADU98749;
 XX
 XX 24-FEB-2005 (first entry)
 DE Borrelia burgdorferi antigenic polypeptide seqid 8.
 XX
 XX antibacterial; vaccine; immune stimulation; immunity; antigen;
 KW DNA library.
 XX
 XX Borrelia burgdorferi.
 OS
 XX WO2004103269-A2.
 PN
 XX
 PD 02-DEC-2004.
 XX
 PF 17-OCT-2003; 2003WO-US033056.
 XX
 PR 18-OCT-2002; 2002US-0419401P.
 XX
 PA (MACR-) MACROGENICS INC.
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX
 XX Sykes KF, Hale KS, Johnston SA;
 PI
 XX
 DR WPI; 2004-834155/82.
 DR N-PSDB; ADU98749.
 XX
 XX Immunizing a subject against Borrelia burgdorferi infection comprises
 PT providing to the subject at least one Borrelia antigen or its fragment.
 XX
 XX Claim 27; SEQ ID NO 8; 121pp; English.
 XX
 XX The invention describes a method of immunizing a subject comprising
 CC providing to the subject at least one Borrelia antigen or its fragment to
 CC induce an immune response. Also described are: an isolated polynucleotide
 CC comprising a sequence having at least 17 contiguous nucleotides in common
 CC with a sequence not given in the specification; an isolated polypeptide
 CC having at least 5 consecutive amino acids of the sequence not given in
 CC the specification; a vaccine composition comprising at least one Borrelia
 CC antigen or at least one polynucleotide encoding a Borrelia antigen;
 CC screening for at least one test polypeptide or test polynucleotide
 CC encoding a polypeptide for an ability to produce an immune response;

CC preparing a vaccine; vaccinating a subject; treating a subject infected
 CC with a pathogen; passive immunization; vaccination; detecting Herpesvirus
 CC ; and detecting antibodies to a herpesvirus. The method is useful in
 CC immunizing a subject against Borrelia burgdorferi infection. This is the
 CC amino acids sequence of a Borrelia burgdorferi polypeptide that may be
 CC useful in immunizing a subject against infection by a member of the
 CC Borrelia species.
 XX
 SQ Sequence 290 AA;
 Query Match 100.0%; Score 1527; DB 8; Length 290;
 Best Local Similarity 100.0%; Pred. No. 5e-137;
 Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 MYKFLFFIIFMFLSCDEKSKSNLKVIGYVNWGGETAATNVLKVFEKMGYNAEIFS 60
 1 MYKFLFFIIFMFLSCDEKSKSNLKVIGYVNWGGETAATNVLKVFEKMGYNAEIFS 60
 61 VTTISIMYOYLAGSKIDGTSSWVPTADKFFYEKLTKEFVDLGNAYEGTIQGFVPSYVPI 120
 61 VTTISIMYOYLAGSKIDGTSSWVPTADKFFYEKLTKEFVDLGNAYEGTIQGFVPSYVPI 120
 121 SSISLKGKGDGFKNKMGIDAGAGTQIVTEQALNYGLSKEYELVPSSSVMLASLDSS 180
 121 SSISLKGKGDGFKNKMGIDAGAGTQIVTEQALNYGLSKEYELVPSSSVMLASLDSS 180
 181 IKRNEWILVPLWKPHWAFSRDYIKFLDDPDLIMGGIESVHTLVRLGLENDDFDAYVFDH 240
 181 IKRNEWILVPLWKPHWAFSRDYIKFLDDPDLIMGGIESVHTLVRLGLENDDFDAYVFDH 240
 241 FYWSDDLILPLMDKNDKEPKGEYRNAVEFVEKKEIVKTWPEKYKTLFD 290
 241 FYWSDDLILPLMDKNDKEPKGEYRNAVEFVEKKEIVKTWPEKYKTLFD 290
 RESULT 3
 AAY19955
 ID AAY19955 standard; protein; 275 AA.
 AC AAY19955;
 XX
 XX 19-JUL-1999 (first entry)
 XX
 DE B. burgdorferi antigenic protein, t810.aa.
 XX
 XX Antigenic protein; vaccine; Lyme disease; infection; detection.
 KW
 XX Borrelia burgdorferi.
 OS
 XX WO9859071-A1.
 PN
 XX 30-DEC-1998.
 PD
 XX 18-JUN-1998; 98WO-US012718.
 PF
 XX 20-JUN-1997; 97US-0050359P.
 PR 22-JUL-1997; 97US-0053344P.
 PR 22-JUL-1997; 97US-0053377P.
 PR 03-SEP-1997; 97US-0057483P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (MEDI-) MEDIMMUNE INC.
 PA
 XX
 XX Choi GH, Erwin AL, Hanson MS, Lathigra R;
 PI
 XX WPI; 1999-189980/16.
 DR N-PSDB; AAX61652.
 DR
 XX New isolated Borrelia burgdorferi nucleic acids - used to develop
 PT products for the diagnosis, prevention and treatment of diseases caused
 PT by Borrelia, particularly Lyme disease.
 XX
 PS Claim 12; Page 134; 275pp; English.

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:21:03 ; Search time 64.8074 Seconds
(without alignments)
1966.133 Million cell updates/sec

Title: US-10-688-058-8
Perfect score: 1527
Sequence: 1 MYKLFPIIFMFLSCDEKK.....EKNKEIVKTWPEKYKTLFD 290

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq 21:**

- 1: Geneseqp1980s:**
- 2: Geneseqp1990s:**
- 3: Geneseqp2000s:**
- 4: Geneseqp2001s:**
- 5: Geneseqp2002s:**
- 6: Geneseqp2003as:**
- 7: Geneseqp2003bs:**
- 8: Geneseqp2004s:**
- 9: Geneseqp2005s:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1527	100.0	290	2	AAV19954
2	1527	100.0	290	8	ADU98749
3	1451	95.0	275	2	AAV19955
4	493	32.3	577	7	ADC95590
5	468	30.6	91	8	ADU98747
6	459.5	30.1	309	9	ABM96406
7	459	30.1	569	4	AU34910
8	455	29.8	571	7	ADH87483
9	451	29.5	573	5	ABB54794
10	442.5	29.0	575	5	ABP27731
11	439	28.7	575	5	ABP27730
12	439	28.7	575	5	ABP29826
13	439	28.7	575	8	ADV88795
14	439	28.7	575	8	ADV80048
15	439	28.7	575	8	ADH82185
16	438	28.7	384	7	ABO72983
17	433.5	28.4	320	6	ABD07852
18	416.5	27.3	334	6	ADH07856
19	416.5	27.3	352	6	ADH07858
20	408	26.7	302	6	ADB11550
21	406	26.6	303	6	ADH07978
22	406	26.6	326	6	ADH07980
23	338.5	22.2	273	6	ADH07854
24	315	20.6	349	4	AAU46688

25	315	20.6	349	6	ABM43207
26	294	19.3	417	4	AAU40421
27	294	19.3	417	6	ABM36940
28	279	18.3	300	5	ABB48738
29	269	17.6	356	7	ABO68411
30	262	17.2	296	8	ADG32088
31	246.5	16.1	427	7	ABO73023
32	235	15.4	412	4	AAU33717
33	235	15.4	312	6	ABU15624
34	235	15.4	318	7	ABO68427
35	224	14.7	380	6	ABU21819
36	189	12.4	314	6	ABU39563
37	175	11.5	333	6	ABM69068
38	166	10.9	361	7	ABO84191
39	160.5	10.5	352	7	ABO84186
40	144.5	9.5	338	7	ABO63655
41	118	7.7	158	5	ABP40562
42	118	7.7	158	8	ADH06144
43	118	7.7	352	7	ADF05311
44	113	7.4	165	5	ABP40486
45	113	7.4	165	8	ADS06036

ALIGNMENTS

RESULT 1
AAV19954
ID AAV19954 standard; protein; 290 AA.
XX
AC AAV19954;
XX
DT 19-JUL-1999 (first entry)
XX
DE B. burgdorferi antigenic protein, f810.aa.
XX
KW Antigenic protein; vaccine; Lyme disease; infection; detection.
XX
OS Borrelia burgdorferi.
XX
PN WO9859071-Al.
XX
PD 30-DEC-1998.
XX
PF 18-JUN-1998; 98WO-US012718.
XX
PR 20-JUN-1997; 97US-0050359P.
PR 22-JUL-1997; 97US-0053344P.
PR 22-JUL-1997; 97US-0053377P.
PR 03-SEP-1997; 97US-0057483P.
XX
(HUMA-) HUMAN GENOME SCI INC.
(MEDI-) MEDIMMUNE INC.
Choi GH, Erwin AL, Hanson MS, Lathigra R;
WPI; 1999-189980/16.
DR N-P5DB; AAX61651.
XX
New isolated Borrelia burgdorferi nucleic acids - used to develop products for the diagnosis, prevention and treatment of diseases caused by Borrelia, particularly Lyme disease.
Claim 12; Page 134; 275pp; English.
XX
This sequence represents a Borrelia burgdorferi (Bb) protein of the invention, which is suitable for use in a vaccine. The Bb polypeptides can be used in vaccines for eliciting protective antibodies to members of the Borrelia genus, particularly for the use against Lyme disease in humans and animals. They can be used for preventing or attenuating an infection caused by a member of the Borrelia genus. The products can also be used for detection of members of the Borrelia genus

PI Sykes KF, Hale KS, Johnston SA;
 XX WPI: 2004-834155/82.
 DR N-PSDB; ADU98746.
 XX
 PT Immunizing a subject against *Borrelia burgdorferi* infection comprises
 PT providing to the subject at least one *Borrelia* antigen or its fragment.
 XX
 XX Claim 27; SEQ ID NO 6; 121pp; English.
 XX
 CC The invention describes a method of immunizing a subject comprising
 CC providing to the subject at least one *Borrelia* antigen or its fragment to
 CC induce an immune response. Also described are: an isolated polynucleotide
 CC comprising a sequence having at least 17 contiguous nucleotides in common
 CC with a sequence not given in the specification; an isolated polypeptide
 CC having at least 5 consecutive amino acids of the sequence not given in
 CC the specification; a vaccine composition comprising at least one *Borrelia*
 CC antigen or at least one polynucleotide encoding a *Borrelia* antigen;
 CC screening for at least one test polypeptide or test polynucleotide;
 CC preparing a vaccine; vaccinating a subject; treating a subject infected
 CC with a pathogen; passive immunization; vaccination; detecting Herpesvirus
 CC ; and detecting antibodies to a herpesvirus. The method is useful in
 CC immunizing a subject against *Borrelia burgdorferi* infection. This is the
 CC amino acids sequence of a *Borrelia burgdorferi* polypeptide that may be
 CC useful in immunizing a subject against infection by a member of the
 CC *Borrelia* species.
 XX
 SQ Sequence 91 AA;
 Query Match 30.6%; Score 468; DB 8; Length 91;
 Best Local Similarity 98.9%; Pred. No. 1.3e-36;
 Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 109 IQGFVVPVSVPISSISELKGDKFKNNKMGIDAGAGTQIVTEQALNYGLSKVELVPS 168
 Db 1 LQGFVVPVSVPISSISELKGDKFKNNKMGIDAGAGTQIVTEQALNYGLSKVELVPS 60
 QY 169 SESVMLASLDSSTKRNWILVPLKPHWAFS 199
 Db 61 SESVMLASLDSSTKRNWILVPLKPHWAFS 91
 RESULT 6
 ABM96406
 ID ABM96406 standard; protein; 309 AA.
 XX
 AC ABM96406;
 DT 02-JUN-2005 (first entry)
 XX
 DE M. xanthus protein sequence, seq id 15605.
 XX
 KW Transgenic plant; DNA replication; gene regulation; gene expression.
 XX
 OS Myxococcus xanthus.
 XX
 PN US6833447-B1.
 XX
 PD 21-DEC-2004.
 XX
 PF 10-JUL-2001; 2001US-00902540.
 XX
 PR 10-JUL-2000; 2000US-0217883P.
 XX
 PA (MONS) MONSANTO TECHNOLOGY LLC.
 XX
 PI Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;
 XX
 DR WPI; 2005-028716/03.
 XX
 PT New substantially purified Myxococcus xanthus nucleic acid molecule
 PT encoding a nitrite reductase, useful for determining gene expression,

PT identifying mutations in a gene of interest, and for constructing
 PT mutations in a gene of interest.
 XX
 PS Example 2; SEQ ID NO 15605; 25pp; English.
 XX
 CC The invention relates to a substantially purified nucleic acid molecule
 CC encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
 CC recombinant DNA construct for expression of a nitrite reductase gene in a
 CC plant cell, and a plant cell comprising the recombinant DNA construct.
 CC The nucleic acid is useful for determining gene expression, identifying
 CC mutations in a gene of interest, and for constructing mutations in a gene
 CC of interest. Sequences given in records for SEQ IDs 9692-16825 represent
 CC a group of 7134 *Myxococcus xanthus* proteins and peptides. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from USPTO
 XX
 SQ Sequence 309 AA;
 Query Match 30.1%; Score 459.5; DB 9; Length 309;
 Best Local Similarity 37.7%; Pred. No. 4.6e-35;
 Matches 98; Conservative 50; Mismatches 103; Indels 9; Gaps 4;
 QY 28 VKIGYVNWGGTAATNLKVVFE-RMGYNASIFSVTTSIMYOYLAGSKIDGTVSSWVPTA 86
 Db 54 VRLVYVNWAEGVAMTHLVQAILEDPMGYEVKTTMADVAPVFAASLANGADAFDGLWLPVT 113
 QY 87 DKFYEKLTKEFDLGANYEGTIOGFVVPVSVPISSISELKGDKFKNNKMGIDAGAGT 146
 Db 114 HQSYMERFQGVLDGTNYEDARIGLVVPADLITSIQLNGKAKDLNQSIVGIDSGAGI 173
 QY 147 QIVTEQALNYGLSKVELVPSSESVMLASLDSSTKRNWILVPLKPHWAFSRYDIKFL 206
 Db 174 MTTTEKALAEYKL--DLKLVSSGPMATLAKKRPVVVTGKPKHWKFAEWDLREL 231
 QY 207 DDPDLIMGIGISVHTLVRLGLENDDFDAYYVFDHPYWSDDLILPLMD---KNDKBPGEY 263
 Db 232 DDPKGVYKAKESIHTLTVGLEKDLPDVATLLRNFKLDDQQLGSLMGAIAEAEDAPEKAT 291
 QY 264 RNAVEFVEKKEIKYKTVYPE 283
 Db 292 R---EWVKKNQALVDGWIPK 308
 RESULT 7
 AAU34910
 ID AAU34910 standard; protein; 569 AA.
 XX
 AC AAU34910;
 DT 14-FEB-2002 (first entry)
 XX
 DE Enterococcus faecalis cellular proliferation protein #197.
 XX
 KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.
 XX
 OS Enterococcus faecalis.
 XX
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US009180.
 XX
 PR 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI: 2001-611495/70.
 DR N-PSDB; AAS52769.
 XX
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 XX
 PS Example 3; SEQ ID NO 10503; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes
 CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
 CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
 CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 569 AA;

Query Match 30.1%; Score 459; DB 4; Length 569;
 Best Local Similarity 35.5%; Pred. No. 1.2e-34;
 Matches 93; Conservative 59; Mismatches 102; Indels 8; Gaps 4;
 QY 21 SSKNLKVKIGVNWGGTAATNLKVYFPEKMGYNAEIFSVTTSIMYQYLAGSKIDGTVS 80
 DB 312 TTNDKQISLSVDEWTEVASTHVAEVLKDMGYDKTTPLDNAIMWESVAKGETDAMVG 371
 QY 81 SWVPADKFFYEKLTKEFDLGANYEGTIGQFVPSYVPISSISELKGKDFKKNMIGI 140
 DB 372 AWLPGTHAEQYKQYKDKLDDLGENLKGAKLGIVVPSYMDVDSIEDLSQAGK---KITGI 428
 QY 141 DAGAGTQIVTEQALNYGLSKEYELVPSSESVMLASLDSSIKRNEWILVPLKPHWAFSR 200
 DB 429 EPGAGVVAAREKTEAYPNLKDWSVETSSSGAMTVAGQAIKNEDIVITGWSPHWMAF 488
 QY 201 YDIKFLDDPDLIMGIESVHTLVRLGLENDPDFAYVDFHFW-SDDLILPLMDKNDKEP 259
 DB 489 YDLKYLADPKGTMGGEAHTHARQGLKEDQPEAYKVLDFNFWHTTKDMESVMLEINE--- 545
 QY 260 GKEYRNAV-EFVEKNKEIVKTV 280
 DB 546 GKDPQEAARDWVDSHKQVAEW 567

RESULT 8
 ADH87483
 ID ADH87483 standard; protein; 571 AA.
 XX
 AC ADH87483;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Enterococcus faecalis polypeptide #1963.
 XX
 KW Enterococcus faecalis infection; transcription regulatory element;
 KW antibacterial.
 XX
 OS Enterococcus faecalis.
 XX

PN US6617156-B1.
 XX
 PD 09-SEP-2003.
 XX
 PF 13-AUG-1998; 98US-00134000.
 XX
 PR 15-AUG-1997; 97US-0055778P.
 XX
 PA (DOUC/) DOUCETTE-STAMM L A.
 PA (BUSH/) BUSH D.
 XX
 PI Doucette-Stamm LA, Bush D;
 XX
 DR WPI: 2003-895394/82.
 DR N-PSDB; ADH84078.
 XX
 PT New nucleic acid comprising a sequence encoding an *Enterococcus faecalis*
 PT polypeptide, useful for preparing a composition for diagnosing or
 PT treating *E. faecalis* infection.
 XX
 PS Disclosure; SEQ ID NO 5368; 193pp; English.
 XX
 CC The invention relates to *Enterococcus faecalis* polynucleotides and
 CC polypeptides. The invention also relates to a recombinant expression
 CC vector comprising a polynucleotide operably linked to a transcription
 CC regulatory element, a cell comprising a recombinant vector, a method for
 CC producing an *E. faecalis* polypeptide, an isolated nucleic acid comprising
 CC a sequence not given in the specification, a recombinant vector
 CC comprising the nucleic acid and a cell comprising the recombinant vector.
 CC The polynucleotides can be used to detect the presence of *E. faecalis* in
 CC a sample. The sequences are useful for preparing a composition for
 CC diagnosing or treating *Enterococcus faecalis* infection. This sequence
 CC represents an *E. faecalis* polypeptide of the invention.
 XX
 SQ Sequence 571 AA;

Query Match 29.8%; Score 455; DB 7; Length 571;
 Best Local Similarity 35.1%; Pred. No. 2.9e-34;
 Matches 92; Conservative 60; Mismatches 102; Indels 8; Gaps 4;
 QY 21 SSKNLKVKIGVNWGGTAATNLKVYFPEKMGYNAEIFSVTTSIMYQYLAGSKIDGTVS 80
 DB 314 TTNDKQISLSVDEWTEVASTHVAEVLKDMGYDKTTPLDNAIMWESVAKGETDAMVG 373
 QY 81 SWVPADKFFYEKLTKEFDLGANYEGTIGQFVPSYVPISSISELKGKDFKKNMIGI 140
 DB 374 AWLPGTHAEQYKQYKDKLDDLGENLKGAKLGIVVPSYMDVDSIEDLSQAGK---KITGI 430
 QY 141 DAGAGTQIVTEQALNYGLSKEYELVPSSESVMLASLDSSIKRNEWILVPLKPHWAFSR 200
 DB 431 EPGAGVVAAREKTEAYPNLKDWSVETSSSGAMTVAGQAIKNEDIVITGWSPHWMAF 490
 QY 201 YDIKFLDDPDLIMGIESVHTLVRLGLENDPDFAYVDFHFW-SDDLILPLMDKNDKEP 259
 DB 491 YDLKYLADPKGTMGGEAHTHARQGLKEDQPEAYKVLDFNFWHTTKDMESVMLEINE--- 547
 QY 260 GKEYRNAV-EFVEKNKEIVKTV 280
 DB 548 GKDPQEAARDWVDSHKQVAEW 569

RESULT 9
 ABB54794
 ID ABB54794 standard; protein; 573 AA.
 XX
 AC ABB54794;
 XX
 DT 29-AUG-2003 (revised)
 DT 16-MAY-2002 (first entry)
 XX
 XX Lactococcus lactis protein busAB.
 DE Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
 KW

XX Lactococcus lactis; IL11403.
 XX FR2807446-A1.
 XX 12-OCT-2001.
 XX 11-APR-2000; 2000FR-00004630.
 XX 11-APR-2000; 2000FR-00004630.
 XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
 XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;
 XX WPI; 2002-043418/06.
 XX New nucleotide sequence useful in the identification or Lactococcus
 XX lactis and related species.
 XX Claim 6; SEQ ID NO 1496; 2504pp; French.
 XX The present invention is related to a Lactococcus lactis nucleotide
 XX sequence (ABA90521) and related proteins (AB53300-AB55621). The nucleic
 XX acid sequence is useful in the detection and/or amplification of nucleic
 XX acid sequence, particularly to identify Lactococcus lactis or related
 XX species. The proteins of the invention are useful for the biosynthesis or
 XX biodegradation of a composition of interest. The invention helps research
 XX in lactic bacteria, particularly useful in the production of yogurt and
 XX cheese. Note: The sequence data for this patent is based on equivalent
 XX patent WO200177334 (published 18-OCT-2001) which is available in
 XX electronic format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences. (Updated on 29-AUG-2003 to
 XX standardise OS field)
 XX Sequence 573 AA;
 Query Match 29.5%; Score 451; DB 5; Length 573;
 Best Local Similarity 34.9%; Pred. No. 7e-34; Mismatches 99; Indels 22; Gaps 7;
 Matches 99; Conservative 64; Mismatches 99; Indels 22; Gaps 7;
 QY 6 LFFIIFMFLSCDEKSSKNLKSIVKIGYVNWGGETAATNVLKVVFERKMGYNABIFSVTTSI 65
 Db 302 LALIIGAFSGMSFGKTASD-KKVDLYVMWDSVASINVLTOAKKEHGFVDKTTALDNVAV 360
 QY 66 MYQYLASGKIDGTVSSWFTADKFFYEKLTKEFVD-LGANYEGTQGFVVPSPYVPISSIS 124
 Db 361 AWQTVANGQADGWSAWLENTHKTQWKY-GKSDVLLGPNLKGAKVGFVVPSPYVNVNIE 419
 QY 125 ELKGGDKFKNMIGIDAGAGTQIVTEQALNYGLSKVEYLPSSSEVNLASLDSIKRN 184
 Db 420 DITNQANK--TITGLEPAGVMAASEKTLSYDNLKMWLPSSSGAMTVALGEAIKQH 476
 QY 185 EWILVPLMKPHAFGRYDIKFLDDPDLIMGGIESVHTLVRLGLENDPDAIYVVFHFYV- 243
 Db 477 KDIVITGSPHWMENKYDLKYLADPKGTGTSENINTIVRGLKKENPEAYKVLKFNWT 536
 QY 244 SDDLILPLMD-KNDKEPGKEYNNAVEFEKNIKEIVKTWVPEKYK 286
 Db 537 TKDMEAVMLDIQNGKTP-----EEAAKNWIKDHQK 566
 RESULT 10
 ABP27731
 ID ABP27731 standard; protein; 575 AA.
 XX AC
 XX ABP27731;
 XX 02-JUL-2002 (first entry)
 DT Streptococcus polypeptide SEQ ID NO 4638.
 DE Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX Streptococcus pyogenes.
 XX WO200234771-A2.
 XX 02-MAY-2002.
 XX 29-OCT-2001; 2001WO-GB004789.
 XX 27-OCT-2000; 2000GB-00026333.
 PR 24-NOV-2000; 2000GB-00028727.
 PR 07-MAR-2001; 2001GB-00005640.
 XX (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX Telford J, Masignani V, Margarit Y Rosi, Grandi G, Fraser C;
 PI Tettelin H;
 XX WPI; 2002-352536/38.
 DR N-PSDB; ABN68362.
 XX New Streptococcus protein for the treatment or prevention of infection or
 XX disease caused by Streptococcus bacteria, such as meningitis, and for
 XX detecting a compound that binds to the protein.
 PS Claim 1; Page 3626-3627; 4525pp; English.
 XX The invention relates to a protein (ABP25413-ABP30895) from group B
 XX streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 XX (Streptococcus pyogenes), comprising one of 5493 sequences (Sl), given in
 XX the specification. The proteins have antibacterial and antiinflammatory
 XX activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 XX antibodies that bind (I) are used in the manufacture of medicaments for
 XX the treatment or prevention of infection or disease caused by
 XX Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 XX Nucleic acids encoding (I) are used to detect Streptococcus in a
 XX biological sample. (I) is used to determine whether a compound binds to
 XX (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 XX used as a vaccine or diagnostic composition. The disease caused by
 XX Streptococcus that is prevented or treated may be meningitis. Nucleic
 XX acid encoding (I) may be used to recombinantly produce (I) and may be
 XX used in gene therapy. Antibodies to (I) are used for affinity
 XX chromatography, immunossays, and distinguishing/identifying
 XX Streptococcus proteins
 XX Sequence 575 AA;
 Query Match 29.0%; Score 442.5; DB 5; Length 575;
 Best Local Similarity 35.8%; Pred. No. 4.6e-33;
 Matches 92; Conservative 56; Mismatches 102; Indels 7; Gaps 5;
 QY 26 KSVKIGYVNWGGETAATNVLKVVFERKMGYNABIFSVTTSIMYQYLASGKIDGTVSSWVPT 85
 Db 322 ETVNIAIYQWDSSEVASTHIAEVLKNEGHVTLTFLDNVAVMMQTIVANGNADFSTSAWLPV 381
 QY 86 ADKFYIEKLTKEFVDLGANYEGTQGFVVPSPYV-PISSISELKGKDKFKNMIGIDAGA 144
 Db 382 THGQQYQKYKSLDDLGNLKGTLGLAVPKYMTVDVNSIEDLSKQADQ---KITGEPGA 438
 QY 145 GTQIVTEQALNYGLSKVEYLPSSSEVNLASLDSIKRNEWILVPLMKPHAFGRYDIK 204
 Db 439 GIMAAQKTLKYEYHNLSSWELVAASGTAMTTSLSDOAIKKKDPVVVTAWSPHWMFAKYDLK 498
 QY 205 FLDDPDLIMGGIESVHTLVRLGLENDPDAIYVVFHFYVSD-DLILPLMDKNDKEPGKEV 263
 Db 499 YLKDPKEIFGTSENINTIARKKELPNVYKIIDKFHTQKMEAVMLDIN-KGMSPB- 556
 QY 264 RNAVEFEKNIKEIVKTW 280
 Db 557 AAARKWVEANKSVSSW 573

Db	366	TVANGNADFTTSAMLPKTHGQVFNKYKNSLDDLGLPHVENVKIGLVVFKMNVNSIEELSN	425
Qy	129	KGDFKFKMIGIDAGAGTQIVTEQALNYYGLSKEYELVPSSESVMLASDSSIKENWIL	188
Db	426	QADK---QITGLEPGAGIMKSAKQSLKDYPNLSSWKLLSASTGAMTTTLGKAIKNKQV	482
Qy	189	VPLWKPWAFSRDYDKFLDDPDLIMGGIESVHTLVRLGLENDDFDAYVYVDFHYHSDLLI	248
Db	483	ITGSPHMFYAKYDLKYLKDPKKSFGGSEHINTIARKNLKDKMPKVYKIIDFKFKTKEDM	542
Qy	249	LPLMDKNDK--EPGKEYRNAVFEVKEKNEIKVKTW	280
Db	543	ESIMLMDKMGEPAKA---AQKWIKNHKEVSEW	573
RESULT 12			
ABP29826	ID	ABP29826 standard; protein; 575 AA.	
XX	AC	ABP29826;	
XX	DT	02-JUL-2002 (first entry)	
XX	DE	Streptococcus polypeptide SEQ ID NO 8828.	
XX	DE	Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;	
KW	KW	group A streptococcus; Streptococcus pyogenes; antibacterial;	
KW	KW	antiflammatory; infection; vaccine; meningitis; gene therapy.	
XX	OS	Streptococcus agalactiae.	
XX	PN	WO200234771-A2.	
XX	PD	02-MAY-2002.	
XX	PF	29-OCT-2001; 2001WO-GB004789.	
XX	PR	27-OCT-2000; 2000GB-00026333.	
PR	PR	24-NOV-2000; 2000GB-00028727.	
PR	PR	07-MAR-2001; 2001GB-00005640.	
XX	XX	(CHIR-) CHIRON SPA.	
PA	PA	(GENO-) INST GENOMIC RES.	
PI	PI	Telford J, Massignani V, Margarit Y RosI, Grandi G, Fraser C;	
PI	PI	Tettelin H;	
XX	XX	WPI; 2002-352536/38.	
DR	DR	N-PSDB; ABN70457.	
XX	XX	New Streptococcus protein for the treatment or prevention of infection or	
PT	PT	disease caused by Streptococcus bacteria, such as meningitis, and for	
PT	PT	detecting a compound that binds to the protein.	
XX	PS	Claim 1; Page 4000; 4525pp; English.	
XX	CC	The invention relates to a protein (ABP25413-ABP30895) from group B	
CC	CC	streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS	
CC	CC	(Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in	
CC	CC	the specification. The proteins have antibacterial and antiinflammatory	
CC	CC	activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and	
CC	CC	antibodies that bind (I) are used in the manufacture of medicaments for	
CC	CC	the treatment or prevention of infection or disease caused by	
CC	CC	Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.	
CC	CC	Nucleic acids encoding (I) are used to detect Streptococcus in a	
CC	CC	biological sample. (I) is used to determine whether a compound binds to	
CC	CC	(I). A composition comprising (I) or a nucleic acid encoding (I), may be	
CC	CC	used as a vaccine or diagnostic composition. The disease caused by	
CC	CC	Streptococcus that is prevented or treated may be meningitis. Nucleic	
CC	CC	acid encoding (I) may be used to recombinantly produce (I) and may be	
CC	CC	used in gene therapy. Antibodies to (I) are used for affinity	
CC	CC	chromatography, immunoassays, and distinguishing/identifying	
CC	CC	Streptococcus proteins	
XX	SQ	Sequence 575 AA;	
Query Match 28.7%; Score 439; DB 5; Length 575;			
Best Local Similarity 33.9%; Pred. No. 9.8e-33;			
Matches 93; Conservative 55; Mismatches 114; Indels 12; Gaps 4;			
Qy	9	IIFMFLSCDEKSSKNLKVGIYVNWGGETAATNLVKVFERMGYNABIFSVTTSIMYQ	68
Db	310	IIVNMTSGNEARGKQK---VKIAYVQWDSEVASTNVIAEVLKSGYDVELTPLDNVAVWQ	365
Qy	69	YIASKGIDGTVSSWPTADKFFYEKLTFRVDLGNAYETIOGFVVPSTVPISSISLKG	128

CC nucleotide sequences (I; ADV78860-ADV78998 and ADV83341-ADV85476) and
 CC novel polypeptides (II; ADV78999-ADV81203 and ADV81205-ADV83340). The
 CC nucleotide sequences encode polypeptides of S. agalactiae involved in the
 CC synthesis of amino acids, cell membranes, intermediate (central)
 CC metabolism, energetic metabolism, fatty acid and phospholipid metabolism,
 CC nucleotide metabolism including purines, pyrimidines and/or nucleosides,
 CC regulatory functions, replication, transcription, translation, protein
 CC transport, adaptation to atypical conditions, sensitivity to medicines
 CC and/or analogues, functions related to transposons, biosynthesis of
 CC cofactors, prosthetic groups and transporters, cell membrane proteins and
 CC cellular machinery. (I) are useful for the detection and/or amplification
 CC of nucleic acids. Pharmaceutical composition comprising (I) or (II) are
 CC useful for treatment of a bacterial S. agalactiae infection. The complete
 CC genome of Streptococcus agalactiae is given in ADV81204. Note: The
 CC present patent is an equivalent for the basic patent FR2824074A1, which
 CC contains only 2344 sequences.

XX Sequence 575 AA;

Query Match 28.7%; Score 439; DB 8; Length 575;
 Best Local Similarity 33.9%; Pred. No. 9.8e-33;
 Matches 93; Conservative 55; Mismatches 114; Indels 12; Gaps 4;

QY 9 IIFMFLSCDEKSSKNLKSXKIGYNNWGGTAAATNVLVKVFEMGYNABIFSVTTSIMYQ 68
 DB 310 IVVNTSGNEAGQK----VKIAYVQWDSSEVASTNVIAEVLKSKGYDVELTFLDNVMMQ 365
 QY 69 YLASGKIDGTSSWVPTADKFEYKELKTFVDLGANYECTIOGVVPSVYVPISSISELKG 128
 DB 366 TVANGNADFTTSAWLPKTHGQYFNKYKNSLDLGHVENVKIGLVVPKMYNVSIEELSN 425
 QY 129 KGDGKFNKMGIDAGAGTQIVTEQALNYYGLSKYELVPSSESVMLASLDSSIKNEWIL 188
 DB 426 QADK---QITGIEPGAGIMKSAKQSLKDYPNLSSWKLSSASTGAMTTTLGKAIKNKQV 482
 QY 189 VPLWKPWFAPSRDYDKFLDDPDLIMGGIESVHTLVRLGLENDPDAIYVDFHYSDDL 248
 DB 483 ITGWSPHWFAKYDLKYLKDPKKSFGGEEHINTIARKNLKDKMPKYKIIDKFKWTKEDM 542
 QY 249 LPLMDKNDK--EPGKEYRNAVEFVEKNKEIVKTW 280
 DB 543 ESIMLMDKMGEPAKA---AQKWKHNKKEVSEW 573

RESULT 15

ADV82185
 ID ADV82185 standard; protein; 575 AA.

AC ADV82185;

XX 24-FEB-2005 (first entry)

XX Streptococcus agalactiae protein, SEQ ID 3326.

XX Antibacterial; vaccine; bacterial infection.

XX Streptococcus agalactiae.

XX WO200292818-A2.

XX 21-NOV-2002.

XX 26-APR-2002; 2002WO-IB0003059.

XX 26-APR-2001; 2001FR-00005642.

XX (INSP) INST PASTEUR.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX Glaser P, Rusniok C, Chevalier F, Frangeul L, Lalioui L;

PI Zouine M, Couve E, Buchrieser C, Poyart C, Trieu-Cuot P, Kunst P;

XX WPI; 2004-101891/11.

XX

PT Genomic nucleotide sequences encoding polypeptides of Streptococcus
 PT agalactiae for the development of vaccines, diagnostic tools, DNA chips
 PT and identification of therapeutic targets.

PS Claim 6; SEQ ID NO 3326; 439pp; French.

XX

CC The present invention relates to novel Streptococcus agalactiae
 CC nucleotide sequences (I; ADV78860-ADV78998 and ADV83341-ADV85476) and
 CC novel polypeptides (II; ADV78999-ADV81203 and ADV81205-ADV83340). The
 CC nucleotide sequences encode polypeptides of S. agalactiae involved in the
 CC synthesis of amino acids, cell membranes, intermediate (central)
 CC metabolism, energetic metabolism, fatty acid and phospholipid metabolism,
 CC nucleotide metabolism including purines, pyrimidines and/or nucleosides,
 CC regulatory functions, replication, transcription, translation, protein
 CC transport, adaptation to atypical conditions, sensitivity to medicines
 CC and/or analogues, functions related to transposons, biosynthesis of
 CC cofactors, prosthetic groups and transporters, cell membrane proteins and
 CC cellular machinery. (I) are useful for the detection and/or amplification
 CC of nucleic acids. Pharmaceutical composition comprising (I) or (II) are
 CC useful for treatment of a bacterial S. agalactiae infection. The complete
 CC genome of Streptococcus agalactiae is given in ADV81204. Note: The
 CC present patent is an equivalent for the basic patent FR2824074A1, which
 CC contains only 2344 sequences.

XX Sequence 575 AA;

Query Match 28.7%; Score 439; DB 8; Length 575;
 Best Local Similarity 33.9%; Pred. No. 9.8e-33;
 Matches 93; Conservative 55; Mismatches 114; Indels 12; Gaps 4;

QY 9 IIFMFLSCDEKSSKNLKSXKIGYNNWGGTAAATNVLVKVFEMGYNABIFSVTTSIMYQ 68
 DB 310 IVVNTSGNEAGQK----VKIAYVQWDSSEVASTNVIAEVLKSKGYDVELTFLDNVMMQ 365
 QY 69 YLASGKIDGTSSWVPTADKFEYKELKTFVDLGANYECTIOGVVPSVYVPISSISELKG 128
 DB 366 TVANGNADFTTSAWLPKTHGQYFNKYKNSLDLGHVENVKIGLVVPKMYNVSIEELSN 425
 QY 129 KGDGKFNKMGIDAGAGTQIVTEQALNYYGLSKYELVPSSESVMLASLDSSIKNEWIL 188
 DB 426 QADK---QITGIEPGAGIMKSAKQSLKDYPNLSSWKLSSASTGAMTTTLGKAIKNKQV 482
 QY 189 VPLWKPWFAPSRDYDKFLDDPDLIMGGIESVHTLVRLGLENDPDAIYVDFHYSDDL 248
 DB 483 ITGWSPHWFAKYDLKYLKDPKKSFGGEEHINTIARKNLKDKMPKYKIIDKFKWTKEDM 542
 QY 249 LPLMDKNDK--EPGKEYRNAVEFVEKNKEIVKTW 280
 DB 543 ESIMLMDKMGEPAKA---AQKWKHNKKEVSEW 573

Search completed: January 24, 2006, 19:44:29

Job time : 65.8074 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:28:54 ; Search time 11.0018 Seconds
(without alignments)
2536.214 Million cell updates/sec

Title: US-10-688-058-8
Perfect score: 1527
Sequence: 1 MYKLFPLFIIFMFLSCDEKK.....EKNKEIVKTVVPEKYKTLFD 290

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1527	100.0	290	2 H70117	glycine betaine, L
2	451	29.5	573	2 C86806	hypothetical prote
3	438	28.7	285	2 H83241	probable glycine b
4	360	23.6	287	2 A33578	glycine betaine/l-
5	316.5	20.7	292	2 A13042	hypothetical prote
6	316.5	20.7	304	2 C98243	hypothetical prote
7	279	18.3	300	2 A11201	glycine betaine AB
8	278	18.2	300	2 A11559	glycine betaine AB
9	269	17.6	312	2 C82974	hypothetical prote
10	252	16.5	293	2 I40537	glycine betaine AB
11	246.5	16.1	307	2 D83642	hypothetical prote
12	240	15.7	312	2 D95891	probable glycine-b
13	235	15.4	312	2 H82972	hypothetical prote
14	226.5	14.8	317	2 AC3128	conserved hypotet
15	226.5	14.8	363	2 F98159	hypothetical prote
16	216	14.1	326	2 AC3307	glycine betaine/l-
17	206.5	13.5	308	2 AE2829	hypothetical prote
18	206.5	13.5	337	2 C97607	hypothetical prote
19	172	11.3	317	2 AH2856	hypothetical prote
20	159.5	10.4	340	2 H83007	hypothetical prote
21	159	10.4	322	2 H83009	probable binding p
22	146	9.6	265	2 E97633	hypothetical prote
23	141.5	9.3	333	2 G95380	probable periplasm
24	138	9.0	333	2 D95968	probable amino aci
25	137.5	9.0	331	2 AC0842	glycine betaine-bi
26	136	8.9	334	2 AG0322	glycine betaine-bi
27	133.5	8.7	426	2 A13053	hypothetical prote
28	133.5	8.7	426	2 D98232	hypothetical prote
29	132	8.6	330	1 BLECGP	glycine betaine/pr

30	132	8.6	330	2 F91071	glycine betaine/pr
31	131.5	8.6	336	2 AG2686	ABC transporter, s
32	131.5	8.6	336	2 D97468	hypothetical prote
33	128	8.4	330	2 A85916	hypothetical prote
34	121	7.9	345	2 AG3186	hypothetical prote
35	112.5	7.4	261	2 AD2747	hypothetical prote
36	112.5	7.4	261	2 C97528	antigenic protein
37	107.5	7.0	345	2 AD3024	hypothetical prote
38	107.5	7.0	359	2 E98260	periplasmic bindin
39	104.5	6.8	319	2 AC3523	aliphatic sulfonat
40	102.5	6.7	210	2 I40272	outer surface prot
41	102	6.7	271	2 AB1292	amino acid ABC tra
42	100.5	6.6	570	2 G88098	protein F18A12.3 [
43	100.5	6.6	663	2 S40308	p100 protein - Lym
44	100	6.5	271	2 A10532	probable lipoprote
45	99.5	6.5	205	2 I40105	outer surface prot

ALIGNMENTS

RESULT 1

H70117

glycine betaine, l-proline ABC transporter, glycine/betaine/L-proline-binding protein (l
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C:Accession: H70117
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kervatage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A>Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: H70117
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-290 <KLE>
A:Cross-references: UNIPROT:O51169; UNIPARC:UPI000005737B; GB:AE001125; GB:AE000783; NII
A:Experimental source: strain B31

Query Match	100.0%	Score 1527;	DB 2;	Length 290;
Best Local Similarity	100.0%	Pred. No. 1.9e-111;	Mismatches 0;	Indels 0;
Matches 290;	Conservative	0;	0;	Gaps 0;
Qy	1	MYKLFPLFIIFMFLSCDEKKSKNLKSVKIGVNMGGTAATNLKVVFKMGYNAEIPS	60	
Db	1	MYKLFPLFIIFMFLSCDEKKSKNLKSVKIGVNMGGTAATNLKVVFKMGYNAEIPS	60	
Qy	61	VTTSIMYQLASGKIDGTVSSWVPTADKFFYEKLTAKFVDLGANYEGTIQGFVWPSYVPI	120	
Db	61	VTTSIMYQLASGKIDGTVSSWVPTADKFFYEKLTAKFVDLGANYEGTIQGFVWPSYVPI	120	
Qy	121	SSISELKKGKDFKFNKMGIDAGAGTQIVTEQALNYGLSKEYELVPSSESVMLASLDS	180	
Db	121	SSISELKKGKDFKFNKMGIDAGAGTQIVTEQALNYGLSKEYELVPSSESVMLASLDS	180	
Qy	181	IKRNEWILVPLWKPHWAFSRDYDIKFLDDPDLIMGGIESVHTLVRLGLENDPDAVYVFDH	240	
Db	181	IKRNEWILVPLWKPHWAFSRDYDIKFLDDPDLIMGGIESVHTLVRLGLENDPDAVYVFDH	240	
Qy	241	FYWSDDLILPLMDKNDKEPGKYEYRANAVFVEKNKEIVKTVWPEKYKTLFD	290	
Db	241	FYWSDDLILPLMDKNDKEPGKYEYRANAVFVEKNKEIVKTVWPEKYKTLFD	290	

RESULT 2

C86806

hypothetical protein busAB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: C86806
R:Boletín, A.; Wincker, P.; Mager, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich

Db 32 KDLGEQINYYTITGIDAGAGIMLATQNAIKDYHLDDNNQLOTSSTAAMTSTLQAKMDKR 91
Qy 186 WTLVPLWKPWFAPSYDYDKFLDDPDLIMGGIESVHTLVRLGLENDPDAYYVDFHYW-S 244
Db 92 PIVVTGWTPHMFTKFDLKLDDPKKRVFGNAENIHTIVKGLKEBKSAKYKVLDNFFWTA 151
Qy 245 DDLILPLMDKNDKBEQKRYNAVEFVEKNKEIVKWTW 280
Db 152 EDMSEVMLENDGVDPBE--AAKKWIKNNPEKVAKM 185
RESULT 9
C2974
hypothetical protein PA5378 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
A:Accession: C82974
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br-
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho-
gen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: C82974
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-312 <STO>
A:Cross-references: UNIPROT:Q9HT16; UNIPARC:UPI00000C5FCF; GB:AE004950; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA5378

Query Match 17.6%; Score 269; DB 2; Length 312;
Best Local Similarity 28.2%; Pred. No. 1.7e-13;
Matches 75; Conservative 57; Mismatches 120; Indels 14; Gaps 6;
Qy 27 SVKIGVYVNGGTAATNVLKLVFVKMGYNABIFSVTTSIMYOYLAGSKIDGVVSSWVPTA 86
Db 29 SVRFADVGVHTDITVTAIVTRQVLESIGYTKVKNLISVPYTRSLANNLDLVLGNWPTM 88
Qy 87 D---RFFYEKLTKFVDLGANYEGTIGQFVFPYSYV---PISSISELKGKDKFKNKMIGI 140
Db 89 TNDIKQYAEKGTVE--TLRANLEGAKYTLAVPQYVYVGGVGLRSPADIAPKSDKLGKNIYGI 146
Qy 141 DAGAGTQIVTEQAL--NYGLSKVELVPSSESVMLASLDSSIKENWILVPLWKPWFAP 198
Db 147 EPGNDGNRVAQSMIDKNAFELGK-FKLVESSBAGMLSQVQRAIRRNQWVVFGLWEPHPMN 205
Qy 199 SYRDIKFLDDPDLIMG--GIBSVHTLVRLGLENDPDAYYVDFHYWSDDLILPLMDKN 255
Db 206 TRFQWKYLEGGDDPFGPNYGGATITVNRKGYAQECANVGQLLKNLSFTLEMNKLMNAV 265
Qy 256 DKEPGKEYRNAVEFVEKNKEIVKWTW 281
Db 266 LNNKKPBEAAKAWLKHDPHEQLDAWL 291

RESULT 10
I40537
glycine betaine ABC transporter (glycine betaine-binding protein) opuAC precursor - Bac-
illus subtilis
C:Species: Bacillus subtilis
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004
A:Accession: I40537; F69669
R:Kempf, B.; Bremer, E.
J. Biol. Chem. 270, 16701-16713, 1995
A:Title: OpuA, an osmotically regulated binding protein-dependent transport system for
betaine
A:Reference number: A57322; MUID:95348093; PMID:7622480
A:Accession: I40537
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-293 <RES>
A:Cross-references: UNIPROT:P46922; UNIPARC:UPI000005FED3; EMBL:U17292; NID:g984802; PID
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter

C.; Bron, S.; Brouillette, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Friz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall-
ier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadai, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpatra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.P.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: F69669
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-293 <KUN>
A:Cross-references: UNIPARC:UPI000005FED3; GB:Z99105; GB:AL009126; NID:g2632457; PIDN:CA
A:Experimental source: strain 168
C:Genetics:
A:Gene: opuAC

Query Match 16.5%; Score 252; DB 2; Length 293;
Best Local Similarity 33.1%; Pred. No. 3.3e-12;
Matches 53; Conservative 34; Mismatches 71; Indels 2; Gaps 2;
Qy 121 SSISELKGKDKFKNKMIGIDAGAGTQIVTBOALNYGLSKVELVPSSESVMLASLDSS 180
Db 23 SENDENASAAEQVNKTIIGIDPGGIMSLTDKAMKDYDLN-DWTLISASSAAMTATLKKS 81
Qy 181 IKRNEWILVPLWKPWFAPSYDYDKFLDDPDLIMGGIESVHTLVRLGLENDPDAYYVDFH 240
Db 82 YDRKKPIIITGWTTHMFWFSRYLKYLDLPDKQSYGSAEIIHTITRKGFSQEPNAAKLLSQ 141
Qy 241 FVWSDLLILPLMDKNDKEPGKEYRNAVEFVEKNKEIVKWTW 280
Db 142 FKWTQDEMGEMIKVE-EGEPKPAKVAAYVNNKHDDQIAEW 180

RESULT 11
D83642
hypothetical protein PA0030 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
A:Accession: D83642
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br-
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho-
gen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: D83642
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-307 <STO>
A:Cross-references: UNIPROT:Q9I7A0; UNIPARC:UPI00000C4EF4; GB:AE004442; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0030

Query Match 16.1%; Score 246.5; DB 2; Length 307;
Best Local Similarity 30.4%; Pred. No. 9.5e-12;
Matches 63; Conservative 42; Mismatches 91; Indels 11; Gaps 7;
Qy 28 VKIGVYVNGGTAATNVLKV-VFEKMGYNABIFSVTTSIMYOYLAGSKIDGVVSSWVPTA 86
Db 30 VRLADPGW-SDIATVNATAAFLLSLGQVKIDILSVPIIYVGLRDGQVDAFLGGWMP-A 87
Qy 87 DKFYFEKLV--KTKFVVDLGANYEGTIGQFVFPYSYV---PISSISELKGKDKFKNKMIGID 141
Db 88 HQDYHDKFVASSQVERLGNLDGTRFTLAVPRYVWDAGVHRFPEDLAAQQRNFKLYGIG 147


```
QY 142 AGAGTQIVTEQAL--NYGLSKEYELVPSSSEVMLASLDSSIKRNEWILVPLWKPHWAFS 199
      ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 154 PCAPANQNIQKIIDANDFGL-KGWQLVESGEQAQVERAGRDKKAVVFLAWAPHPMNE 212
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 200 RYDIKFLDDPDLIMG---GIESVHTLVRLGLENDDEFDAYVFDHFYW-----SDDLILPLM 252
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 213 KIDIAVLSGGDAYFGPNFGAEVYTLARTGWQPQCNAATLFPNLKFDIGMENSIMGSL 272
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 253 DKNDKEPGKEYRNAVEFEKNKEIVKTVV 281
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 273 GGEDSKAA-----AAAWLKANPKALDGL 296
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

RESULT 15
F98159
hypothetical protein AGR_L_470 [imported] - Agrobacterium tumefaciens (strain C58, Cerec
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: F98159
A.: Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: F98159
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-363 <KUR>
A:Cross-references: UNIPROT:Q8U709; UNIPARC:UPI00000D264E; GB:AE007870; PIDN:AAK88800.1;
C:Genetics:
A:Gene: AGR_L_470
A:Map position: linear chromosome

Query Match 14.8%; Score 226.5; DB 2; Length 363;
Best Local Similarity 23.8%; Pred. No. 4.3e-10;
Matches 64; Conservative 63; Mismatches 123; Indels 19; Gaps 7;

QY 26 KSVKIGYVNWGGETAATNVLKVVEKMGYNASIFSVTTSIMYQYLASGKIDGTSSWVPT 85
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 80 RTIRLSDFGWTITATNGVASYVLLDALGYEADVKTLSVPIGYQAMKNGEIDVFLGNWMPA 139
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 86 ADKFYEEKLTKFVD-LGANYEGTIQGFVVPISYVP---ISSISELKGKDKFKNMIGID 141
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 140 QKSFVDDLNAKAAAEVMTKMLEGAKFTLAVPAYMAEKGVKDFADLAKHADEFESKIYIE 199
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 142 AGAGTQIVTEQAL--NYGLSKEYELVPSSSEVMLASLDSSIKRNEWILVPLWKPHWAFS 199
      ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 200 PCAPANQNIQKIIDANDFGL-KGWQLVESGEQAQVERAGRDKKAVVFLAWAPHPMNE 258
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 200 RYDIKFLDDPDLIMG---GIESVHTLVRLGLENDDEFDAYVFDHFYW-----SDDLILPLM 252
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 259 KIDIAVLSGGDAYFGPNFGAEVYTLARTGWQPQCNAATLFPNLKFDIGMENSIMGSL 318
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 253 DKNDKEPGKEYRNAVEFEKNKEIVKTVV 281
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 319 GGEDSKAA-----AAAWLKANPKALDGL 342
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
```

Search completed: January 24, 2006, 19:59:02
Job time : 12.0018 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:30:59 ; Search time 17.8779 Seconds
(without alignments)
1341.095 Million cell updates/sec

Title: US-10-688-058-8
Perfect score: 1527
Sequence: 1 MYKFLFLFIIFMFLSCDEKK.....EKNKEIVKTVWPEKTKTLFD 290

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5 COMB.pcp.*
2: /cgn2_6/ptodata/1/1aa/6 COMB.pcp.*
3: /cgn2_6/ptodata/1/1aa/H COMB.pcp.*
4: /cgn2_6/ptodata/1/1aa/PCTUS COMB.pcp.*
5: /cgn2_6/ptodata/1/1aa/RE COMB.pcp.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1527	100.0	290	US-09-830-230A-361	Sequence 361, App
2	1451	95.0	275	US-09-830-230A-362	Sequence 362, App
3	493	32.3	577	US-09-107-532A-5217	Sequence 5217, App
4	459.5	30.1	309	US-09-902-540-15605	Sequence 15605, A
5	455	29.8	571	US-09-134-000C-5368	Sequence 5368, App
6	438	28.7	384	US-09-252-991A-17129	Sequence 21729, A
7	269	17.6	356	US-09-252-991A-17157	Sequence 17157, A
8	246.5	16.1	427	US-09-252-991A-21769	Sequence 21769, A
9	235	15.4	318	US-09-252-991A-17173	Sequence 17173, A
10	166	10.9	361	US-09-252-991A-32937	Sequence 32937, A
11	160.5	10.5	352	US-09-252-991A-32932	Sequence 32932, A
12	144.5	9.5	338	US-09-489-039A-10172	Sequence 10172, A
13	118	7.7	158	US-09-134-001C-5407	Sequence 5407, App
14	118	7.7	352	US-09-543-681A-5596	Sequence 5596, App
15	113	7.4	165	US-09-134-001C-5331	Sequence 5331, App
16	108.5	7.1	266	US-08-961-083-14	Sequence 14, Appl
17	108.5	7.1	266	US-09-536-784-14	Sequence 14, Appl
18	108.5	7.1	266	US-09-765-271-14	Sequence 14, Appl
19	108.5	7.1	266	US-09-765-272A-14	Sequence 14, Appl
20	106.5	7.0	285	US-09-489-039A-10641	Sequence 10641, A
21	100.5	6.6	337	US-09-328-352-6980	Sequence 6980, App
22	100.5	6.6	663	US-09-196-293-5	Sequence 5, Appl
23	100.5	6.6	663	US-08-209-603B-5	Sequence 5, Appl
24	100.5	6.6	663	US-09-711-546-5	Sequence 5, Appl
25	100.5	6.6	663	US-10-289-795-5	Sequence 5, Appl
26	98.5	6.5	663	US-08-235-836C-78	Sequence 78, Appl
27	97.5	6.4	823	US-09-107-532A-5667	Sequence 5667, App

28 95 6.2 221 2 US-09-107-433-2796 Sequence 2796, App
29 95 6.2 372 2 US-09-710-279-182 Sequence 182, App
30 95 6.2 391 2 US-09-134-001C-3952 Sequence 3952, App
31 95 6.2 1751 2 US-09-136-574A-44 Sequence 44, Appl
32 94 6.2 990 2 US-10-101-464A-814 Sequence 814, App
33 93.5 6.1 773 1 US-08-966-389-4 Sequence 4, Appl
34 93.5 6.1 773 1 US-09-103-509-4 Sequence 4, Appl
35 93.5 6.1 773 1 US-09-102-644-4 Sequence 4, Appl
36 93.5 6.1 773 1 US-09-218-032-4 Sequence 4, Appl
37 93.5 6.1 1426 2 US-09-136-574A-43 Sequence 43, Appl
38 92.5 6.1 700 2 US-08-235-836C-74 Sequence 74, Appl
39 92 6.0 1881 2 US-09-233-086-3 Sequence 3, Appl
40 91.5 6.0 441 2 US-09-328-352-5114 Sequence 5114, App
41 91.5 6.0 663 2 US-08-235-836C-70 Sequence 70, Appl
42 91 6.0 339 2 US-09-107-532A-5639 Sequence 5639, App
43 90.5 5.9 287 2 US-09-107-532A-4780 Sequence 4780, App
44 90.5 5.9 317 2 US-09-252-991A-31329 Sequence 31329, A
45 90 5.9 279 2 US-09-543-681A-6534 Sequence 6534, App

ALIGNMENTS

RESULT 1

US-09-830-230A-361
; Sequence 361, Application US/09830230A
; Patent No. 6902893
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481US
; CURRENT APPLICATION NUMBER: US/09/830,230A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: PCT/US98/12718
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/057,483
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 361
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-830-230A-361

Query Match 100.0%; Score 1527; DB 2; Length 290;
Best Local Similarity 100.0%; Pred. No. 8.6e-150; Indels 0; Gaps 0;
Matches 290; Conservative 0; Mismatches 0;

QY 1 MYKFLFLFIIFMFLSCDEKKSKNLKSVKIGYVNMGGTAAATNLVKWPEKMGYNAEIFS 60
Db 1 MYKFLFLFIIFMFLSCDEKKSKNLKSVKIGYVNMGGTAAATNLVKWPEKMGYNAEIFS 60

QY 61 VTTISIMYOYLASGKIDGTIVSSWVPADKFFYEKLTAKTFVDLGCANYEGTIGQFVPSYVPI 120
Db 61 VTTISIMYOYLASGKIDGTIVSSWVPADKFFYEKLTAKTFVDLGCANYEGTIGQFVPSYVPI 120

QY 121 SSISLKGKGDKFKNMKGIDAGAGTQVTEQALNYGLSKYELVPSSVLMASLDS 180
Db 121 SSISLKGKGDKFKNMKGIDAGAGTQVTEQALNYGLSKYELVPSSVLMASLDS 180

QY 181 IKRNEWILVPLWKPHWAFSRDYDIKFLDDPDLIMGGIESVHTLVRLGLENDDFDAYVYFDH 240
Db 181 IKRNEWILVPLWKPHWAFSRDYDIKFLDDPDLIMGGIESVHTLVRLGLENDDFDAYVYFDH 240

QY 241 FWSDDLILPLMKNDKPGKRYNAVFVEKNKIVKTVWPEKTKTLFD 290
Db 241 FWSDDLILPLMKNDKPGKRYNAVFVEKNKIVKTVWPEKTKTLFD 290

Db 241 FYWSDLLPLMDKNDKEGKEYNNAVEFEKKEIKVKTWVPEKYKTLFD 290

RESULT 2

US-09-830-230A-362
; Sequence 362, Application US/09830230A
; Patent No. 6902893

GENERAL INFORMATION:

; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481US
; CURRENT APPLICATION NUMBER: US/09/830,230A
; CURRENT FILING DATE: 2001-09-27

; PRIOR APPLICATION NUMBER: PCT/US98/12718

; PRIOR FILING DATE: 1998-06-18

; PRIOR APPLICATION NUMBER: 60/057,483

; PRIOR FILING DATE: 1997-09-03

; PRIOR APPLICATION NUMBER: 60/053,344

; PRIOR FILING DATE: 1997-07-22

; PRIOR APPLICATION NUMBER: 60/053,377

; PRIOR FILING DATE: 1997-07-22

; PRIOR APPLICATION NUMBER: 60/050,359

; PRIOR FILING DATE: 1997-06-20

; NUMBER OF SEQ ID NOS: 756

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 362

; LENGTH: 275

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-830-230A-362

Query Match 95.0%; Score 1451; DB 2; Length 275;

Best Local Similarity 100.0%; Pred. No. 6e-142;

Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 CDEKSSKNLKSVKIGYNWGGTAAATNVLKVFVKMGYNABIFSVTTSIMYQYLASGKI 75

Db 1 CDEKSSKNLKSVKIGYNWGGTAAATNVLKVFVKMGYNABIFSVTTSIMYQYLASGKI 60

QY 76 DGTSSWVPTADKYFEKLEKTKFVLDGANYEGTIQGFVVPSPVPISSISELKGKDGKFN 135

Db 61 DGTSSWVPTADKYFEKLEKTKFVLDGANYEGTIQGFVVPSPVPISSISELKGKDGKFN 120

QY 136 KWIGDAGAGTQVTEQALNYGLSKEYELVPSSSVMLASLDSISKRNWILVPLMKPH 195

Db 121 KWIGDAGAGTQVTEQALNYGLSKEYELVPSSSVMLASLDSISKRNWILVPLMKPH 180

QY 196 WAFSRYDIKFLDDPDLIMGIESVHTLVRLGLENDDFDAYYVDFHFWSDDLILPLMDKN 255

Db 181 WAFSRYDIKFLDDPDLIMGIESVHTLVRLGLENDDFDAYYVDFHFWSDDLILPLMDKN 240

QY 256 DKEPGKEYNNAVEFEKKEIKVKTWVPEKYKTLFD 290

Db 241 DKEPGKEYNNAVEFEKKEIKVKTWVPEKYKTLFD 275

RESULT 3

US-09-107-532A-5217

; Sequence 5217, Application US/09107532A

; Patent No. 6583275

GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 5217:

SEQUENCE CHARACTERISTICS:

LENGTH: 577 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...577

SEQUENCE DESCRIPTION: SEQ ID NO: 5217:

US-09-107-532A-5217

Query Match 32.3%; Score 493; DB 2; Length 577;

Best Local Similarity 37.0%; Pred. No. 3.7e-42;

Matches 97; Conservative 65; Mismatches 92; Indels 8; Gaps 4;

QY 21 SSKNLKSVKIGYNWGGTAAATNVLKVFVKMGYNABIFSVTTSIMYQYLASGKIDGTVS 80

Db 320 SAKETKRLNSYVEMDEVASTNVVGEVLKQMGYDVTMTPLDINSIMKSVSNGESDAMVS 379

QY 81 SWVPADKFYFEKLEKTKFVLDGANYEGTIQGFVVPSPVPISSISELKGKDFKKNMIGI 140

Db 380 AMLPETHGSOYAQYKQVEDIGANLTGAKVGLAVPAYMDVNSIDELTQAGK--XIGI 436

QY 141 DAGAGTQVTEQALNYGLSKEYELVPSSSVMLASLDSISKRNWILVPLMKPHAFSR 200

Db 437 EPGAGVWTAANTIQKYNLNDKWKVETSSSGAMTVALGQAIKKHSPVIVVTGTPHWPFAK 496

QY 201 YDIKFLDDPDLIMGIESVHTLVRLGLENDDFDAYYVDFHFWSD-DLILPLMD-KNDKE 258

Db 497 YDLKYLEDPENGMGSEEQIHTMVRKGLKBDQPEAYKVLNDPHWSEKDMKVMLEINNGKD 556

QY 259 PGKEYRNAVEFEKKEIKVKTW 280

Db 557 P---QQAADWKIKENQELVESW 575

RESULT 4

US-09-902-540-15605

; Sequence 15605, Application US/09902540

; Patent No. 6833447

GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:59:25 ; Search time 4.64138 Seconds
(without alignments)
633.203 Million cell updates/sec

Title: US-10-688-058-8
Perfect score: 1527
Sequence: 1 MYKLFPIIFLSCDEKK.....EKNKEIVKTWPEKYKTLFD 290

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 70609 seqs, 10134256 residues

Total number of hits satisfying chosen parameters: 70609

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08 NEW PUB.pap.*
- 2: /cgn2_6/ptodata/1/pubpaa/US06 NEW PUB.pap.*
- 3: /cgn2_6/ptodata/1/pubpaa/US07 NEW PUB.pap.*
- 4: /cgn2_6/ptodata/1/pubpaa/US09 NEW PUB.pap.*
- 5: /cgn2_6/ptodata/1/pubpaa/US10 NEW PUB.pap.*
- 6: /cgn2_6/ptodata/1/pubpaa/US11 NEW PUB.pap.*
- 7: /cgn2_6/ptodata/1/pubpaa/US12 NEW PUB.pap.*
- 8: /cgn2_6/ptodata/1/pubpaa/US60 NEW PUB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	262	17.2	296	US-10-510-386-58	Sequence 58, Appl
2	98.5	6.5	663	US-11-196-475-78	Sequence 78, Appl
3	95	6.2	372	US-10-793-626-182	Sequence 182, Appl
4	92.5	6.1	700	US-11-196-475-74	Sequence 74, Appl
5	91.5	6.0	348	US-11-129-143-72	Sequence 72, Appl
6	91.5	6.0	663	US-11-196-475-70	Sequence 70, Appl
7	89	5.8	693	US-11-196-475-68	Sequence 68, Appl
8	86	5.6	568	US-10-793-626-2482	Sequence 2482, Ap
9	86	5.6	1145	US-10-793-626-1432	Sequence 1432, Ap
10	85.5	5.6	708	US-11-196-475-76	Sequence 76, Appl
11	85	5.6	585	US-11-074-176-190	Sequence 190, Appl
12	84.5	5.5	687	US-11-074-176-260	Sequence 260, Appl
13	82	5.4	429	US-10-793-626-3174	Sequence 3174, Ap
14	82	5.4	693	US-11-196-475-72	Sequence 72, Appl
15	81.5	5.3	700	US-11-196-475-66	Sequence 66, Appl
16	81.5	5.3	2516	US-10-647-956A-2	Sequence 2, Appl
17	81	5.3	271	US-11-052-554A-225	Sequence 225, Appl
18	81	5.3	567	US-10-485-517-216	Sequence 216, Appl
19	81	5.3	743	US-10-485-517-351	Sequence 351, Appl
20	81	5.3	817	US-10-793-626-2948	Sequence 2948, Ap
21	81	5.3	877	US-10-485-517-200	Sequence 200, Appl
22	81	5.3	2710	US-11-051-453-41	Sequence 41, Appl
23	80.5	5.3	255	US-10-485-517-360	Sequence 360, Appl
24	80.5	5.3	392	US-10-793-626-194	Sequence 194, Appl
25	79.5	5.2	175	US-10-793-626-1928	Sequence 1928, Ap

26	79.5	5.2	767	6	US-10-467-962B-91	Sequence 91, Appl
27	79.5	5.2	1889	7	US-11-102-476-46	Sequence 46, Appl
28	79	5.2	387	7	US-11-074-176-232	Sequence 232, App
29	77.5	5.1	513	6	US-10-453-372-442	Sequence 442, App
30	77.5	5.1	535	6	US-10-453-372-436	Sequence 436, App
31	77	5.0	503	7	US-11-019-711-104	Sequence 104, App
32	77	5.0	730	6	US-10-453-372-444	Sequence 444, App
33	77	5.0	737	6	US-10-453-372-450	Sequence 450, App
34	77	5.0	889	7	US-11-038-284-15	Sequence 15, Appl
35	76.5	5.0	1140	6	US-10-858-730-208	Sequence 208, App
36	76	5.0	342	7	US-11-156-084-183	Sequence 183, App
37	76	5.0	411	7	US-11-024-959-330	Sequence 330, App
38	76	5.0	857	7	US-11-052-554A-218	Sequence 218, App
39	76	5.0	1735	6	US-10-495-083-3	Sequence 3, Appli
40	75.5	4.9	272	6	US-10-793-626-2290	Sequence 2290, Ap
41	75.5	4.9	318	7	US-11-194-246-303	Sequence 303, App
42	75.5	4.9	402	6	US-10-485-517-422	Sequence 422, App
43	75.5	4.9	945	7	US-11-024-959-375	Sequence 375, App
44	75	4.9	207	7	US-11-196-475-36	Sequence 36, Appl
45	75	4.9	382	7	US-11-179-411-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-10-510-386-58
; Sequence 58, Application US/10510386
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510.386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 58
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-58

Query Match	17.2%	Score 262;	DB 6;	Length 296;
Best Local Similarity	35.2%	Pred. No. 1.7e-15;		
Matches	57;	Conservative 35;	Mismatches 64;	Indels 6; Gaps 3;
Qy	121	SSISELKGKDKFKNMIGIDAGAGTQIVTQALNYGLSKEYELVPSSEVMLASLDSS	180	
Db	24	SNNENASVGDQVNYKITGIDPGAGIMNATDQALXDYDLSK-WTVTSGSSSANTAAKKA	82	
Qy	181	IKRNEWILVPLKHPWAFSRVDIKELDQDLIMGIESVHTLVRLGLENDPDFDAYVFDH	240	
Db	83	YDKDPPIITGTWPHMFAKYDLKYLKDKPGSGYDAEEIHTVTRKGFDDHFGANKLLSQ	142	
Qy	241	FVMSDDLl--LPLMDKNDKEPGKEYRNAVEFVEKNKEIVKTV	280	
Db	143	FSWTEDDMGVNLAVQEGKKPEEA---AADFVKHQDLVKKW	181	

RESULT 2
US-11-196-475-78
; Sequence 78, Application US/11196475
; Publication NO. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Comes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.

```
; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 663
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
; US-11-196-475-78

Query Match          6.5%; Score 98.5; DB 7; Length 663;
Best Local Similarity 19.9%; Pred. No. 0.53;
Matches 68; Conservative 58; Mismatches 97; Indels 119; Gaps 16;

QY 1 MYKFLFLFIIFMF---LSCDE--KKSKNKLKSVKGYVNWGGTAAATNVLVKVFVKMGYN 55
Db 4 MLLIFSFLLVFLNGFLNAREVDKELKDFVNMDFVNYKGYPDSTNTYEQI---VGIG 60
QY 56 AEIFSFTTSMYQYLASGKIDGTSSVWPTADKFYVEKLTKEFVD-----LGA 103
Db 61 -----EFLARPLINSNS-----SYGKVFVNRFDIDQDKKASVDIFSIGS 102
QY 104 -----NYEGTIQGFVVPSYVPISSISELKGKDKFKNMIGIDAGAGTQIVTEQALN 155
Db 103 KSELDLSILNRLITGLYMKSPDYERSSAELAKATTIYNAVYRGD-----LD 150
QY 156 YGLSKEYELVPSSESVMLASLDSISKRNW-----ILVPLMKPHWAFSRDYIKFLDDPD 210
Db 151 YY---KEFYIASLKLTKENAGLSRVYSQWAGKTQIFLPLK-----N 191
QY 211 LMGGIES---VHTLV-----RLGLEND-----DFDAYVYV-FDHF 241
Db 192 ILSGNVEIDIDSLVTDKVAALLSENGSVNFARDITDIQETHKADQDKIDIELDNF 251
QY 242 YWSDDLILPLMDK-----NDKEPKGYRNAVEFVEKNKE 275
Db 252 HESDSNITETIENLRDLEKATDDEHKKIESQVDAKKQKE 293

RESULT 3
US-10-793-626-182
; Sequence 182, Application US/10793626
; Publication No. US2005055478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 182
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
```

```
US-10-793-626-182

Query Match          6.2%; Score 95; DB 6; Length 372;
Best Local Similarity 19.7%; Pred. No. 0.49;
Matches 54; Conservative 37; Mismatches 75; Indels 108; Gaps 12;

QY 39 TAAATNVLVKVFVKMGYNAEIFSFTTSMYQYLASGKIDGTSSVW-PTADKFYVEKLTKE 97
Db 92 TFTPVPMSVLTEQFNINIPATVMTDYRMH-----KNWITPYSQRYVATKDTK 140
QY 98 --FVDLGANYEGTIQGFVVPSYVPISSISELKGKDKFKNM-----IGIDAGAGTQ 147
Db 141 DDFIEAG-----VPASYIKVTGIP-----IADKFEESIDKEWLSQQHLDPSKPTI 186
QY 148 IVTEQALNYGLSKEYELV-----PSSSVMLASLDSISKRNWILVPLMKPHWAFS 199
Db 187 LMSAGA---FGVSKGFQYMINNILEKSPNSQVWVICGRSKELKRS----- 228
QY 200 RYDIKFLDDPDLIM-----GGIESVHTLVRLGLENDDFDAYVYV 238
Db 229 -LKAKFKDNPSVILIGYTNHNMWASSQLMTKPGGITISEGLSR----- 273
QY 239 DHFYWSDDLILPLMDKNDKEPKGYRNAVEFVEK 272
Db 274 -----CIPMTIFLN-PAPQLEENAYYFESK 297

RESULT 4
US-11-196-475-74
; Sequence 74, Application US/11196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
; US-11-196-475-74

Query Match          6.1%; Score 92.5; DB 7; Length 700;
Best Local Similarity 19.6%; Pred. No. 1.9;
Matches 67; Conservative 58; Mismatches 98; Indels 119; Gaps 16;

QY 1 MYKFLFLFIIFMF---LSCDE--KKSKNKLKSVKGYVNWGGTAAATNVLVKVFVKMGYN 55
Db 4 MLLIFSFLLVFLNGFLNAREVDKELKDFVNMDFVNYKGYPDSTNTYEQI---VGIG 60
QY 56 AEIFSFTTSMYQYLASGKIDGTSSVWPTADKFYVEKLTKEFVD-----LGA 103
Db 61 -----EFLARPLINSNS-----SYGKVFVNRFDIDQDKKASVDIFSIGS 102
QY 104 -----NYEGTIQGFVVPSYVPISSISELKGKDKFKNMIGIDAGAGTQIVTEQALN 155
```


Db 103 KSELDILNRLITCYLMKSFYERSSAELIAKAITIYNAVVRGD-----LD 150
QY 156 YTGKSEYELVPSSEVMASLDSSIKRNEW-----ILVPLMKPHWAFSRDYDKFLDDPD 210
Db 151 YI---KEFYIASLSLTAKENAGLSRVYSQWAGTKQIFPLKK-----N 191
QY 211 LTMGGIES---VHTLV-----RLGLEND-----DPDAYVYV-FPHF 241
Db 192 ILSGNVEDSIDLSLVTDKVVAALLSENGSVNFARDITDIQETHKADQDKIDIELONI 251
QY 242 YMSDDLILFLMDK-----NDKEPGKEYRNAVEFVEKNKE 275
Db 252 HESDSNITETIENLRDLEKATDEEHKKIESQVDAKKQKE 293

RESULT 5

US-11-129-143-72
; Sequence 72, Application US/11129143
; Publication No. US20050266518A1
; GENERAL INFORMATION:
; APPLICANT: BERRY, Alan
; APPLICANT: BRETZEL, Werner
; APPLICANT: HUMBELIN, Markus
; APPLICANT: LOPEZ-ULIBARRI, Rual
; APPLICANT: MAYER, Anne F.
; APPLICANT: YELISEEV, Alexei A.
; TITLE OF INVENTION: IMPROVED ISOPRENOID PRODUCTION
; FILE REFERENCE: C38435/121966
; CURRENT APPLICATION NUMBER: US/11/129,143
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 72
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Thermoplasma acidophilum
US-11-129-143-72

Query Match 6.0%; Score 91.5; DB 7; Length 348;

Best Local Similarity 24.6%; Pred. No. 0.89;

Matches 68; Conservative 29; Mismatches 100; Indels 79; Gaps 14;

QY 23 KULKSVKIGYVWGGETAATNV---LKVPFKMGYNAEIFSTTSIMYOYLASGIDGTV 79
Db 52 KKLKPFMISSMTGGAEIAKRNRLAVALAERFGIMGVGSMRAIV-----DRSTIEDTY 106
QY 80 S-----SWVP-----TADK-----FYEKLKTKFVDLGANYVSGTIOGFVV 114
Db 107 SVINESHVPLKNTANGAPOLVRQDKDQVNSNRDIAYIYDLIKADFLAVHFN---LQEMVQ 163
QY 115 P-----SYVPISISSELKKG-----GDKFKNRMIG--IDAG-----AGTQIVTEQA 153
Db 164 PGSDRNSKGVDRIDKDLGSGFNIIAKETSGSGFSRRRTAELIIDAGVKAIKIEVSGVGTTPAA 223
QY 154 LNYGLSKSEYEL-----VPSSESVMLAS-----LDSSIKRNEWTL-----VPIW 192
Db 224 VEYTRARKENKLEKMRIGETFWNWGIPSPASVYVCSDLAPVIGSGGLRGLDLAKAIAMG 283
QY 193 KPHWAFSRVDIKFLD--DPDLIMGIESVHTLVRLGL 227
Db 284 ATAGGFARSLKQADTDPEMLMKNIELIQREFRVAL 319

RESULT 6

US-11-196-475-70
; Sequence 70, Application US/11196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: Recombinant Constructs of Borrelia

; TITLE OF INVENTION: Burgdorferi
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 663
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-11-196-475-70

Query Match 6.0%; Score 91.5; DB 7; Length 663;

Best Local Similarity 20.6%; Pred. No. 2.1;

Matches 68; Conservative 56; Mismatches 111; Indels 95; Gaps 15;

QY 1 MYKLFLEFIIFMF---LSCDE--KKSSKNLKSVKIGYVNWGGETAATNVLKVPFKMGYN 55
Db 4 MLLPSFELVFLNGPPLNAREVDKEKLDKDFNMDLFEVNYKGPYDSTDTYEQIVGIGBFL 63
QY 56 AEIPSVTTSIMYOYLASGIDGTVSSVPTADKYKFKYKFDLGA-----NYEG 107
Db 64 ARPLNNSNSSY--GKY--FVNRFDIDQDK---KASVDIFSIGSKSELDLSILNLR 114
QY 108 TIQGFVPSYVPISISSELKKGDKFKNMKIGIDAGAGTQIVTEQALNYYGLSEYELVP 167
Db 115 ILTYLMKSFYERSSAELIAKAITIYNAVVRGD-----LDYY---KEFYIEA 159
QY 168 SSESVMASLDSSIKRNEW-----ILVPLMKPHWAFSRDYKFLDDPDILIMGIES---V 219
Db 160 SLKSLTKENAGLSRVYSQWAGTKQIFPLKK-----NILSGNVEDSIDI 203
QY 220 HTLV-----RLGLEND-----DPDAYVYV-FDHFVWSDLLPLMD 253
Db 204 DSVLVTDKVVAALLSENGSVNFARDITDIQETHKADQDKIDIELDNFHSNITETIE 263
QY 254 K-----NDKEPGKEYRNAVEFVEKNKE 275
Db 264 NLRDLEKATDEEHKKIESQVDAKKQKE 293

RESULT 7

US-11-196-475-68
; Sequence 68, Application US/11196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; PRIOR FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18

```

; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-11-196-475-68

Query Match          5.6%; Score 89; DB 7; Length 693;
Best Local Similarity 20.0%; Pred. No. 3.7; Indels 62; Gaps 12;
Matches 63; Conservative 54; Mismatches 136;

Qy 1 MYKLFFFIIFMFLSCD-----EKKSSKNLKSVKIGYVNWGGETAATNVLKVPFEK 51
Db 1 MKKLLIFPFL--ISLNGPLNSRVDKELKDFVNMOLDFVNYKGPYDSTNTYIEQIVGI 59

Qy 52 MGVNAE--IPSVTTSIMYOYLAGSKIDGTVSSWVPTADKFYYEKLTKTFVDLGA----- 103
Db 60 GEPLARPLNSNSIYY-----GKY--FINRFIDQDK-----KASVDVFSIGSRSQLDS 108

Qy 104 --NYEGTIOGFV----VPSYVPISSISE-----LKGKDGKFRKNMIGIDAGAGTQIVTQALN 155
Db 60 GEPLARPLNSNSIYY-----GKY--FINRFIDQDK-----KASVDVFSIGSRSQLDS 108

Qy 109 ILMRLITGLTKSDYERSAELIAKVIITHNAVYRGDLNYYKEVYIEAALKSLTKEN 168
Db 158 --GLSKEYEYELVPSSSVMLASLSSIKRNEWILVPLKPHWAFSRVDIKFL-----DDP 209

Qy 158 --GLSKEYEYELVPSSSVMLASLSSIKRNEWILVPLKPHWAFSRVDIKFL-----DDP 209
Db 169 AGLSRVYSQWAGTQIFILPKNILSGKVESDIDISLTDVKVVAALLSENAEAGVNFARD 228

Qy 209 PDLMGGIESVHTLVRLGLENDPDAFYVDFHYMSDDLLILPLMDK-----NDKEPG 260
Db 229 ITDIQGETHKA-----DQDKID--IELDNVHKSDSNITETIENLRDQLEKATDEEHR 278

Qy 261 KEYRNAVEFVKNEIKVKTWVPEKYKTLFD 290
Db 279 KEIESQVDAKKKQKE 293

RESULT 8
US-10-793-626-2482
; Sequence 2482, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: F034800S
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2482
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-2482

Query Match          5.6%; Score 86; DB 6; Length 568;
Best Local Similarity 19.6%; Pred. No. 5.1;
Matches 70; Conservative 59; Mismatches 124; Indels 104; Gaps 15;

Qy 8 FIIFMFLSCDEKSSKNLKSVKIGYVNWGGETAATNVLKVPFEKMGVNAEIPSVT----- 62
Db 68 FLVFM---DQHRSTEDYVNRITQIFITDAQYIYIQEMQSILDTLQINEEQFEAYIQFO 124

Qy 63 -----TSIMYOYLAGSKIDGTVSSWVPTAD-----KFYYEKLTKTFVDLGA 103
Db 125 QNQEVSAAEQALLNDSNQLFNLYKQKRDILDNAYIMTYDMRESLRNLYESMATDF----- 179

RESULT 10
US-11-196-475-76
; Sequence 76, Application US/11196475

```

```
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Soleccki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; TITLE OF INVENTION: Burgdorferi
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-11-196-475-76

Query Match      5.6%; Score 85.5; DB 7; Length 708;
Best Local Similarity 19.2%; Pred. No. 7.5;
Matches 64; Conservative 57; Mismatches 111; Indels 101; Gaps 15;

Qy      1 MYKFLPLFIIMF---LSCD--EKSSKNLKSVKIGYVNWGGETAATNVLVKVFPEKMGYN 55
Db      4 MLLIESFLFIINGPLNARKVDKELKDFVNMDFEFVNYKGPYSTNTYEQI---VGIG 60

Qy      56 AEIPSVTTSIMVOYLASGKIDGTSSVWPTADKFYVEKLTCTKFDVLGANYEGTIOGFVVP 115
Db      61 -----EFLARPLTNSNS-----SYGKYFNFID-DQDKASVDVFSIS 101

Qy      116 SYVPISSSELKGGKDFKNRMIGIDAGAGTQI-----VTRQALNYGLSKEYELV 166
Db      102 SKSELDLILNRILTYGIKSFYDRSSAELIAKVIITYNAYVRGDLQY---KGFYIE 158

Qy      167 PESSEVMSLSDSSIKRNEW-----ILVPLWPKHAFPSRYDIKFLDDPDLINGGIES--- 218
Db      159 PALKSLTKENAGLSRVYSQWAGKTQIFIPKK-----DILSGNIESDID 202

Qy      219 VHTLV-----RLGL-----ENDDFDAYVVDHFWSDLLILP 250
Db      203 IDSLVTVDKVIALLSENAAGVNFARDITDIOGETHAKDQDKIDT--ELDNTHESDSNITE 260

Qy      251 LMDK-----NDKEPGKEYRNAVEFVEKNKE 275
Db      261 TIENLRDQEKATDREHKKEIESQVDAKKKEKE 293

RESULT 11
US-11-074-176-190
; Sequence 190, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Perill, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US/11/074,176
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 260
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-260

Query Match      5.5%; Score 84.5; DB 7; Length 687;
Best Local Similarity 19.7%; Pred. No. 8.8;
Matches 58; Conservative 45; Mismatches 95; Indels 97; Gaps 14;

Qy      85 TADKFYVEKLTCTKFDVLGANYEG-----TIQGFVVPVSV-----PI 120
Db      100 TTDDIYFEEELKGTAYVYVHVQKEGKASDILLGMSEIKAMFTPTKMRWDSNDFBVRPI 159

Qy      121 SSISELKGKDKFKNMKMGIDAG-----AGTQIVTEQALNYGLSKEYELVPSSE-- 170
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 190
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-190

Query Match      5.6%; Score 85; DB 7; Length 585;
Best Local Similarity 18.6%; Pred. No. 6.4;
Matches 58; Conservative 53; Mismatches 136; Indels 64; Gaps 10;

Qy      21 SSKNLKSVKIGYVNWGGETAATNVLK-----VVFERKMGYNAEIPSVTTSIMYQYLA 71
Db      99 TDKGAATIKI---NQAKTATINIKONVWSDGKPVTAKEYAYEIIANKATKSORVTS 155

Qy      72 S-----GKIDGTSSVWPTADKFYVEKLTCTKFDVLGANYEGTIOGFVVPVSV 119
Db      156 SLADLVGLEEYHEGKSNTISGIEMPDGNGRKRKVLHFKTKMKGMTSG--NGYIMGTAAP 213

Qy      120 ISSISELKGK---GDKFKNMKMGIDAGAGTQIVTEQALNYGLSKEYELVPSSESVMLA 175
Db      214 YHYLKDVKFKDLMSDDKIRKKLPFGAYKLOKVRGQSVTWVPNEYYKGGPKLKITAS 273

Qy      176 -----SLDSSIKRNEWILVPLWPKHAFPSRYDIKFLDDPDLINGGIESVHTLVRLGLEND 230
Db      274 VINPNSVAQSISKNKFDVISVNSQW-----QNVK-----GTKGVNPIAQVPLSV- 318

Qy      231 DFDAYVVDHFWSDDLLEPLMDKNDKBPGEKGYRNAVEFVEKNKEIVK-----T 279
Db      319 ---SYLGFVKGVKDAQRGENVVWPKSKMNNKALRQAIAYGMNVDOVKRYTYGLSFRVPT 375

Qy      280 WPEKYKTLFD 290
Db      376 LIPKQFGDYFD 386

RESULT 12
US-11-074-176-260
; Sequence 260, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Perill, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 260
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-260

Query Match      5.5%; Score 84.5; DB 7; Length 687;
Best Local Similarity 19.7%; Pred. No. 8.8;
Matches 58; Conservative 45; Mismatches 95; Indels 97; Gaps 14;

Qy      85 TADKFYVEKLTCTKFDVLGANYEG-----TIQGFVVPVSV-----PI 120
Db      100 TTDDIYFEEELKGTAYVYVHVQKEGKASDILLGMSEIKAMFTPTKMRWDSNDFBVRPI 159

Qy      121 SSISELKGKDKFKNMKMGIDAG-----AGTQIVTEQALNYGLSKEYELVPSSE-- 170
```

Db 160 HMLVSLFG-NEVIPVKILDTAGRTKTEGRFLGDSVVLNADDDYEDALKDQVYVIANAEER 218
QY 171 -SVMLASLDSSIKRNEWILVP-----LWKPWFASRYDIKFLDDPLIM----- 213
Db 219 KDMIVNQMDLVKENHWQKPDRLLEEVYLVVEYPTVFAGSFDEKYLNPIDEVILITSMK 278
QY 214 -----GGTESVHTLVRLGLENDFF-----DAYVVEDFHFWYS 244
Db 279 DNQRFVEVDENGLKLNHFIARVNG--NKDYLDNVISGNEKVLVARLDDA-----QFFYD 331
QY 245 DDLIILPLMDKNDKPEKRYNAV-----EFVEKNKEI-----VKTW-VPEKYKTLFD 290
Db 332 EDRKYPFLSHFVRLKDVSHDKIGSMAEKVQVRMIGDYLAQRWNLPENVVKDFD 386

RESULT 13

US-10-793-626-3174
; Sequence 3174, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3174
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-3174

Query Match 5.4%; Score 82; DB 6; Length 429;
Best Local Similarity 21.2%; Pred. No. 7.7;
Matches 65; Conservative 42; Mismatches 111; Indels 88; Gaps 15;
QY 6 LFFIIFMLSCDEKSSKNLKSVMKIGYVNWGGETAATNVLKVFEE-KMGYNAEIFSVTIS 64
Db 125 LSLVADLFVDTDD-----TVLLPETHNNGN-----YKLVFSTRGAHINTYSI--- 166
QY 65 IMYQVLASGKIDGTSSWVPTADKFYIEKLTKFVDLGANYEGTIQGFVVPSPVP----- 119
Db 167 ----PDDSGHF--TTSSELVKTLEKYYKDK-----VIIILNYPNPTG-----YTPNKKREV 210
QY 120 ----ISSISLKGKGFKNMIGIDAGAGTQIVTEQALNYYGLSKEYELV----- 166
Db 211 NTIVNAIEBLANK-----GTVKVVTVVDDAYYGLF--YEEVYQOCSIFTALT 253
QY 167 -PSSSEVMLASLDSSIKRNEWILVPLKPHWAFSRVDIKFLDDPDILMGIGIES-VHTLVLR 224
Db 254 QVKSNNLLPVRIDGATK--EFF-----SWGFRVGFMTFGIDHETLKNALSAKVKGLIR 304
QY 225 LGLENDDFDAYVVFHFWYSDDLIILPLMDKNDKPEKRYNAVFEVVEKNKEIVKTVWPEK 284
Db 305 SNISSPLPSQSAIKH-----VLKYEQFDKEIDQINILKERVEYTKQVVDNKNYAK 357
QY 285 YKTLFD 290
Db 358 YWQAYD 363

RESULT 14

US-11-196-475-72
; Sequence 72, Application US/11196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:

; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-11-196-475-72

Query Match 5.4%; Score 82; DB 7; Length 693;
Best Local Similarity 21.0%; Pred. No. 15;
Matches 69; Conservative 53; Mismatches 117; Indels 90; Gaps 15;
QY 1 MYKLFLEFIIFMF-----LSCDE--KKSSKNLKSVMKIGYVNWGGETAATNVLKVF--- 49
Db 1 MKKMLLIFFSLISLNGPFLNAREVDKEKLDKFVNMDFEYNYKGPYSTNTYEQIVGIG 60
QY 50 ERMGYNAEIFSVTISIMYQVLASGKIDGTSSWVPTADKFYIEKLTKF-VDLGANYEET 108
Db 61 EFLARPLNFNSNSYGYKYFINRFIDDDQKK--ASVDVF---SISKSQDLSILNLRRI 115
QY 109 IQGFVVPSPVPISSISELKGKGFKNMIGIDAGAGTQIVTEQALNYYGLSKEYELVPS 168
Db 116 LTGYLIKSFDEYSSAELIAKVIITHNAVYRGD-----LNTY---KIFYIESA 160
QY 169 SESVMLASLDSSIKRNEW-----ILVPLKPHWAFSRVDIKFLDDPDILMGIGIES---VH 220
Db 161 LKSLTKENAGLSRVYSQWAGTKQIFIPUKK-----NILSGKIESDIDID 204
QY 221 TLV-----RLGLEND-----DFDAYVY-FDFHFWSDDLILPLMDK 254
Db 205 SLVTDKVVAAALLSENEAGVNFARDITDIQETHKADQDKIDELDNVHESDSNITETIEN 264
QY 255 -----NDKEPGKEYRNAVEFVEKNKE 275
Db 265 LRDLQEKATDBEHRKEIESQVDAKKKQKE 293

RESULT 15

US-11-196-475-66
; Sequence 66, Application US/11196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836

Search completed: January 24, 2006, 20:55:35
Job time : 5.64138 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:57:15 ; Search time 50.9692 Seconds
(without alignments)
2377.329 Million cell updates/sec

Title: US-10-688-058-8
Perfect score: 1527
Sequence: 1 MYKLFPIIFMFLSCDEKK.....BKNKEIVKTVPEKYKTLFD 290

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.Main:
1: /cgm2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap.*
2: /cgm2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap.*
3: /cgm2_6/ptodata/1/pubpaa/US09_PUBCOMB.pap.*
4: /cgm2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pap.*
5: /cgm2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pap.*
6: /cgm2_6/ptodata/1/pubpaa/US11_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1527	100.0	290	5	US-10-688-058-8
2	1527	100.0	290	5	US-10-994-726-361
3	1451	95.0	275	5	US-10-994-726-362
4	468	30.6	91	5	US-10-688-058-6
5	459	30.1	569	3	US-09-815-242-10503
6	442.5	29.0	575	5	US-10-474-792-100
7	433.5	28.4	320	5	US-10-501-282-1792
8	416.5	27.3	334	5	US-10-501-282-1796
9	416.5	27.3	352	5	US-10-501-282-1798
10	408	26.7	302	5	US-10-501-282-5510
11	406	26.6	303	5	US-10-501-282-1918
12	406	26.6	326	5	US-10-501-282-1920
13	370.5	24.3	871	4	US-10-156-761-14244
14	338.5	22.2	273	5	US-10-501-282-1794
15	235	15.4	312	3	US-09-815-242-5213
16	235	15.4	312	4	US-10-282-122A-43548
17	224	14.7	380	4	US-10-282-122A-49743
18	189	12.4	314	4	US-10-282-122A-57487
19	118	7.7	158	4	US-10-724-972A-5439
20	113	7.4	165	4	US-10-724-972A-5331
21	112.5	7.4	324	4	US-10-156-761-10967
22	108.5	7.1	266	3	US-09-765-272-14
23	108.5	7.1	266	6	US-11-106-649-14
24	106	6.9	333	4	US-10-282-122A-5335
25	101	6.6	1842	4	US-10-243-552-847
26	101	6.6	1842	4	US-10-243-552-887
27	100.5	6.6	570	4	US-10-369-493-5353

28	100.5	6.6	663	4	US-10-289-795-5	Sequence 5, Appli
29	100.5	6.6	663	5	US-10-762-665-5	Sequence 5, Appli
30	100	6.5	271	4	US-10-282-122A-76195	Sequence 76195, A
31	99	6.5	1184	4	US-10-282-122A-53497	Sequence 53497, A
32	98.5	6.5	663	4	US-10-369-100-78	Sequence 78, Appl
33	97.5	6.4	270	4	US-10-282-122A-52838	Sequence 52838, A
34	97.5	6.4	274	3	US-09-971-536-58	Sequence 58, Appl
35	97	6.4	265	5	US-10-514-054-2	Sequence 2, Appli
36	97	6.4	348	4	US-10-282-122A-47221	Sequence 47221, A
37	97	6.4	437	4	US-10-282-122A-61263	Sequence 61263, A
38	96	6.3	565	5	US-10-732-923-1913	Sequence 1913, Ap
39	96	6.3	656	4	US-10-424-599-173577	Sequence 173577,
40	96	6.3	843	4	US-10-282-122A-53395	Sequence 53395, A
41	95	6.2	221	5	US-10-617-320-2796	Sequence 2796, Ap
42	95	6.2	391	4	US-10-724-972A-5743	Sequence 5743, Ap
43	94.5	6.2	536	4	US-10-369-493-6866	Sequence 6866, Ap
44	94.5	6.2	727	4	US-10-437-963-103577	Sequence 103577,
45	94	6.2	271	4	US-10-282-122A-55826	Sequence 55826, A

ALIGNMENTS

RESULT 1
US-10-688-058-8
; Sequence 8, Application US/10688058
; Publication No. US20050058661A1
; GENERAL INFORMATION:
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: HALE, KATHERINE S.
; APPLICANT: JOHNSTON, STEPHEN A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF THE GENUS
; TITLE OF INVENTION: BORRELIA
; FILE REFERENCE: MCRO:00305
; CURRENT APPLICATION NUMBER: US/10/688,058
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 8
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-688-058-8

Query Match	100.0%;	Score 1527;	DB 5;	Length 290;
Best Local Similarity	100.0%;	Pred. No. 2.2e-134;		
Matches 290;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MYKLFPIIFMFLSCDEKKSSKNLKVIGYVNWGGTAATNVLKVVFERKMGYNAEIFS	60	
Db	1	MYKLFPIIFMFLSCDEKKSSKNLKVIGYVNWGGTAATNVLKVVFERKMGYNAEIFS	60	
QY	61	VTTSIMQYLAGKIDGTSSVWPTADKFFYEKLTKFVDLCANVEGTIOGFVPSYVPI	120	
Db	61	VTTSIMQYLAGKIDGTSSVWPTADKFFYEKLTKFVDLCANVEGTIOGFVPSYVPI	120	
QY	121	SSISELKGKDKFKNMIGIDAGAGTQIVTQALNYYGLSKYEYLVPSSESVMLASLDSS	180	
Db	121	SSISELKGKDKFKNMIGIDAGAGTQIVTQALNYYGLSKYEYLVPSSESVMLASLDSS	180	
QY	181	IKRNEWILVPLWKPHWAFPSRYDIKFLDDPDLIMGGIESVHTLVRLGLENDPFDAYVFDH	240	
Db	181	IKRNEWILVPLWKPHWAFPSRYDIKFLDDPDLIMGGIESVHTLVRLGLENDPFDAYVFDH	240	
QY	241	FYWSDDLILPLMDKNDKEPGKEYNVAFVEKNEIKVKTWPEKYKTLFD	290	
Db	241	FYWSDDLILPLMDKNDKEPGKEYNVAFVEKNEIKVKTWPEKYKTLFD	290	

RESULT 2
US-10-994-726-361
; Sequence 361, Application US/10994726

```
; Publication No. US20050147999A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481D1
; CURRENT APPLICATION NUMBER: US/10/994,726
; CURRENT FILING DATE: 2004-11-23
; PRIOR APPLICATION NUMBER: 09/830,230
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: PCT/US98/12718
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/057,483
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 361
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
; US-10-994-726-361

Query Match 100.0%; Score 1527; DB 5; Length 290;
Best Local Similarity 100.0%; Pred. No. 2.2e-134;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYKFLFFIIFWLSGDEKSSKNLKVIGYVNWGGETAATNVLKVPFVKMGYNALFIS 60
Db 1 MYKFLFFIIFWLSGDEKSSKNLKVIGYVNWGGETAATNVLKVPFVKMGYNALFIS 60

QY 61 VTTISIMYQYLASGIDGTSSWVPTADKFYIEKLTKFVDLGANYEGTIQGFVWPSYVPI 120
Db 61 VTTISIMYQYLASGIDGTSSWVPTADKFYIEKLTKFVDLGANYEGTIQGFVWPSYVPI 120

QY 121 SSISELKGKDKPKFKNMIGIDAGAGTQIVTEQALNYGLSKYELVPSSSEVNLASLDSS 180
Db 121 SSISELKGKDKPKFKNMIGIDAGAGTQIVTEQALNYGLSKYELVPSSSEVNLASLDSS 180

QY 181 IKRNEWILVPLKPHWAFSRDYDKFLDDPDLIMGGIESVHTLVRLGLENDDFDAYVYFDH 240
Db 181 IKRNEWILVPLKPHWAFSRDYDKFLDDPDLIMGGIESVHTLVRLGLENDDFDAYVYFDH 240

QY 241 FYWSDDLILPLMDKNDKEPGKEYNNAVEFVEKKNKEIVKTWVPEKYKTLFD 290
Db 241 FYWSDDLILPLMDKNDKEPGKEYNNAVEFVEKKNKEIVKTWVPEKYKTLFD 290

RESULT 3
US-10-994-726-362
; Sequence 362, Application US/10994726
; Publication No. US20050147999A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481D1
; CURRENT APPLICATION NUMBER: US/10/994,726
; CURRENT FILING DATE: 2004-11-23
; PRIOR APPLICATION NUMBER: PCT/US98/12718
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/057,483
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359

; Publication No. US20050147999A1
; GENERAL INFORMATION:
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: HALE, KATHERINE S.
; APPLICANT: JOHNSTON, STEPHEN A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF THE GENUS
; TITLE OF INVENTION: BORRELIA
; FILE REFERENCE: MCRO:003US
; CURRENT APPLICATION NUMBER: US/10/688,058
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
; US-10-688-058-6

Query Match 30.6%; Score 468; DB 5; Length 91;
Best Local Similarity 98.9%; Pred. No. 6.9e-36;
Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 109 IQGFVVVPSVPISSISSELKGGKDKPKFKNMIGIDAGAGTQIVTEQALNYGLSKYELVPS 168
Db 1 IQGFVVVPSVPISSISSELKGGKDKPKFKNMIGIDAGAGTQIVTEQALNYGLSKYELVPS 168

QY 169 SESVNLASLDSSISIKRNEWILVPLWKPHWAFS 199
Db 61 SESVNLASLDSSISIKRNEWILVPLWKPHWAFS 91

RESULT 5
US-09-815-242-10503
; Sequence 10503, Application US/09815242
; Patent No. US20020061569A1
```


GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carz, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 1410
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10503
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10503

Query Match 30.1%; Score 459; DB 3; Length 569;
Best Local Similarity 35.5%; Pred. No. 6e-34;
Matches 93; Conservative 59; Mismatches 102; Indels 8; Gaps 4;
QY 21 SKNLKSVKIGYVNGGETAATNLKVPKMGYNABIFSVTTSIMYQYVLAAGKIDGTGVS 80
DB 312 TTNDKQISLSYVEMDEVASTHVAEVLKMGYDVKTTPLDNAIMWESVAKGETDAMVG 371
QY 81 SWPTADKPYEKLTKFVDIGANVEGTIGQFVVPSPVPISSISELKGKDFKKNMIGI 140
DB 372 AMLPGTHAEQYKQYKLDLGNLKGAKLGIWPSYMDVDSIEDLSQAGK---KITGI 428
QY 141 DAGAGTQIVTEQALNYGLSKEYELVPSSSEVMLASLDSSIKRNEWILVPLKPHWAFSR 200
DB 429 EFGAGVVAAEKTKAYPNLKDWSETSSGAMTVALGQAIKKNEDIVITGWSPHMFAK 488
QY 201 YDIKFLDPLIMGGIESVHTLVRGLNDDDFDAYVDFHFW-SDDLILPLMDKNDKEP 259
DB 489 YDLKYLADPKGTMGGEAHTHARQGLKEDQPEAYKVLNDFHFWTTKDWESVMLEINE--- 545
QY 260 GKEYRNAV-EFVEKKNKEIVKTW 280
DB 546 GKDPQEAARDWVDHSHKQVAEW 567

RESULT 6
US-10-474-792-100
; Sequence 100, Application US/10474792
; Publication No. US20040236072A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Stephen
; APPLICANT: Zagursky, Robert
; APPLICANT: Nickbarg, Elliot
; APPLICANT: Winter, Lourie
; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
; FILE REFERENCE: AM 100399
; CURRENT APPLICATION NUMBER: US/10/474,792

; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 674
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 100
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-474-792-100
Query Match 29.0%; Score 442.5; DB 5; Length 575;
Best Local Similarity 35.8%; Pred. No. 2.1e-32;
Matches 92; Conservative 56; Mismatches 102; Indels 7; Gaps 5;
QY 26 KSVKIGYVNGGETAATNLKVPKMGYNABIFSVTTSIMYQYVLAAGKIDGTGVSVPVT 85
DB 322 ETVNIAYQWDSEVASTHVAEVLKMGYHVTLTPLDNVMMQTVANGNADPSTSAWLPV 381
QY 86 ADKPYEKLTKFVDIGANVEGTIGQFVVPSPV-PISISELKGKDFKKNMIGIDAGA 144
DB 382 THGQYQYKSKLDLGNLKGAKLGIWPSYMDVDSIEDLSQADQ---KITGIEPCA 438
QY 145 GTQIVTEQALNYGLSKEYELVPSSSEVMLASLDSSIKRNEWILVPLKPHWAFSRYDIK 204
DB 439 GIMAAQTKLKEYHNLSSWELVAASTGAMTSLDQAIKKKOPIVVTAWSPHMMFAKYDIK 498
QY 205 FLDDPDLIMGGIESVHTLVRGLNDDDFDAYVDFHFWSD-DLILPLMDKNDKEPVGKEY 263
DB 499 YLKPKEIFGSTENINTIARKLKGKELPNVYKIIDKFHTQKDMEAVALDIN-KGMSPE- 556
QY 264 RNAVEFVEKKNKEIVKTW 280
DB 557 AAARKKWEANKSVSSW 573

RESULT 7
US-10-501-282-1792
; Sequence 1792, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOIOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; TITLE OF INVENTION: POLYPEPTIDE ANTIGENS, IMMUNOGENIC COMPOSITIONS AND USES THEREOF
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; PRIOR FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1792
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Alloiococcus otitidis
US-10-501-282-1792

Query Match 28.4%; Score 433.5; DB 5; Length 320;
Best Local Similarity 33.7%; Pred. No. 6.7e-32;
Matches 93; Conservative 62; Mismatches 110; Indels 11; Gaps 4;
QY 12 MFL-SCDEKSKSNLKS-----KIGVYNGGETAATNLKVPKMGYNABIFSVTTS 64
DB 16 LFLTACQGEDGAADVSVVDGEEIEIAYVWDSEIASTHVGIVGESLGLVDVTLTSLDNA 75
QY 65 IMYQYVLAAGKIDGTGVSVPVTADKPYEKLTKFVDIGANVEGTIGQFVVPSPVPISSIS 124
DB 76 VMQSVANGADAMVSAWLPHTHGDQEAEGYGDQMDHVGNTLEGAKIGLGVPTIMEVDSTA 135

Db 205 PNNLQIDMTYLTGSEVDVFGPDEGAATVSTMTAAGYQAQCNCNVARLLHNLRFSSAQV----- 260

Qy 242 YWSDDLILPLMDKNDKEPGKEYRNNAVEFVEKNKEIVKTVV 281

Db 261 ---SQVMAPILDRT-----QPLDAAQWLKANPEPLKAWL 292

Search completed: January 24, 2006, 20:54:30
Job time : 51.9692 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:27:53 ; Search time 62.3148 Seconds
(without alignments)
3283.383 Million cell updates/sec

Title: US-10-688-058-8

Perfect score: 1527

Sequence: 1 MYKFLFPIIFMFLSCDEKK.....EKNKEIVKWVPEKVKILFD 290

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1527	100.0	290	051169 BORBU	051169 borrelia bu
2	1454.5	95.3	294	Q662L9 BORGA	Q662L9 borrelia ga
3	510	33.4	285	Q729Q3 DESVH	Q729Q3 desulfovibr
4	505	33.1	315	Q8TNY0 METAC	Q8TNY0 methanosarc
5	494.5	32.4	289	Q926L5 LISIN	Q926L5 listeria in
6	478.5	31.3	277	Q8A6X6 BACTIN	Q8A6X6 bacteroides
7	463.5	30.4	306	Q8U4S3 METMA	Q8U4S3 methanosarc
8	462	30.3	282	Q5WJD4 BACSK	Q5WJD4 bacillus cl
9	460.5	30.2	298	Q891M9 CLOTE	Q891M9 clostridium
10	459	30.1	569	Q830X6 ENTFA	Q830X6 enterococcu
11	451	29.5	573	Q9KIF6 9LACT	Q9KIF6 lactococcus
12	451	29.5	573	Q7DAU8 LAGLA	Q7DAU8 lactococcus
13	447	29.3	297	Q8EXX8 OCEIH	Q8EXX8 oceanobacil
14	442.5	29.0	561	Q8K8R4 STRP3	Q8K8R4 streptococc
15	442.5	29.0	575	Q879N8 STRP3	Q879N8 streptococc
16	442.5	29.0	575	Q9A1M9 STRPY	Q9A1M9 streptococc
17	440.5	28.8	575	Q5XE24 STRP6	Q5XE24 streptococc
18	439	28.7	575	Q8XQ4 STRA5	Q8XQ4 streptococc
19	439	28.7	575	Q8E3C3 STRA3	Q8E3C3 streptococc
20	438	28.7	285	Q9HZ04 PSEAE	Q9HZ04 pseudomonas
21	436.5	28.6	575	Q8P2S5 STRP8	Q8P2S5 streptococc
22	433.5	28.4	287	Q4LV54 9BURK	Q4LV54 burkholderi
23	428.5	28.1	573	Q9RQ05 9LACT	Q9RQ05 lactococcus
24	427.5	28.0	283	Q4KFZ7 PSEF5	Q4KFZ7 pseudomonas
25	424.5	27.8	576	Q8DU85 STRMU	Q8DU85 streptococc
26	418	27.4	287	Q63MU2 BURPS	Q63MU2 burkholderi
27	418	27.4	287	Q62DH6 BURMA	Q62DH6 burkholderi
28	416.5	27.3	285	Q63AG1 BACCZ	Q63AG1 bacillus ce
29	416	27.2	284	Q64R03 BACFR	Q64R03 bacteroides
30	415.5	27.2	295	Q6LYW9 METMP	Q6LYW9 methanococc
31	415	27.2	284	Q5LAK8 BACFN	Q5LAK8 bacteroides

```

RESULT 1
ID 051169 BORBU PRELIMINARY; PRT; 290 AA.
AC 051169;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DE Glycine betaine, L-proline ABC transporter, glycine/betaine/L-proline-
binding protein (ProX).
GN OrderedLocusNames=BB0144;
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=139;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
RA Winn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Weidman J.F., Uterback T.R., Watthey L., McDonald L.A.,
RA Artiach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
"Genomic sequence of a Lyme disease spirochaete, Borrelia
burgdorferi."
RL Nature 390:580-586 (1997).
DR EMBL; AE001125; AAC86525.1; -; Genomic_DNA.
DR PIR; H70117; H70117.
DR TIGR; BB0144; -.
GO GO:0005488; P:binding; IEA.
GO GO:0005215; P:transporter activity; IEA.
GO GO:0006810; P:transport; IEA.
DR InterPro; IPR007210; OpuAC_ABC.
DR Pfam; PF04069; OpuAC; 1.
KW Complete proteome.
SQ
SEQUENCE 290 AA; 33263 MW; 43779CF0911FBD5 CRC64;

```

ALIGNMENTS

```

Query Match      100.0%; Score 1527; DB 2; Length 290;
Best Local Similarity 100.0%; Pred. No. 3.2e-108;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYKFLFPIIFMFLSCDEKSKNLSVKIGYVNWGGTAAATNVLKVVFKMGYNNAEITFS 60
DB 1 MYKFLFPIIFMFLSCDEKSKNLSVKIGYVNWGGTAAATNVLKVVFKMGYNNAEITFS 60

QY 61 VTTSMYQYLAGSKIDGTGVSSVWPTADKFFYYEKLKTKFVLDGANYEGTIQGFVPSYVPI 120
DB 61 VTTSMYQYLAGSKIDGTGVSSVWPTADKFFYYEKLKTKFVLDGANYEGTIQGFVPSYVPI 120

QY 121 SSISELKGKDGKFNKMGIGIDAGAGTQIVTQALNYGLSKYEYLVPSSESVMASLDSS 180

```

```
Db 121 SSISELKGKDKFKNMIGIDAGAGTQIVTEQALNYGLSKEYELVPSSESVMLASLDSS 180
Qy 181 IKRNEWILVPLMKPHWAFSRDYDKFLDDPDLIMGGIESVHTLVRLGLENDPDAYVVDH 240
Db 181 IKRNEWILVPLMKPHWAFSRDYDKFLDDPDLIMGGIESVHTLVRLGLENDPDAYVVDH 240
Qy 241 FYWSDDLILPLMDKNDKPGKEYNRAVEFVEKNKEIVKTWVPEKYKTLFD 290
Db 241 FYWSDDLILPLMDKNDKPGKEYNRAVEFVEKNKEIVKTWVPEKYKTLFD 290

RESULT 2
Q66219 BORGA
ID Q66219 BORGA PRELIMINARY; PRT; 294 AA.
AC Q66219
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Glycine betaine, L-proline ABC transporter, binding protein.
GN Name=proX; OrderedLocusNames=BG0144;
OS Borrelia garinii.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=29519;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Pbi;
RA Gloeckner G., Lehmann R., Romualdi A., Pradella S.,
RA Schulte-Speckel U., Wilske B., Stuehn J., Platzer M.;
RT "Comparative analysis of the Borrelia garinii genome.";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000013; RAU07002.1; -; Genomic_DNA.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007210; OpuAC_ABC.
DR Pfam; PF04069; OpuAC; 1.
KW Complete proteome.
SQ SEQUENCE 294 AA; 33680 MW; AFA81026194ABE12 CRC64;

Query Match 95.3%; Score 1454.5; DB 2; Length 294;
Best Local Similarity 93.5%; Pred. No. 1.1e-102;
Matches 272; Conservative 13; Mismatches 5; Indels 1; Gaps 1;

Qy 1 MYKFLFPIIFMF-LSCDEKSKSNLKSXKIGYVNWGGETAATNVLKVFFKMGYNAEIF 59
Db 4 IYKFLTGFIIFLFLSCDEKSKSNLKSXKIGYVNWGGETAATNVLKVFFKMGYNAEIF 63
Qy 60 SVTTSIMYQYLASGKIDGTSSWVPTADKPYEKLTKFVDLGANYEGTIQGFVPSVYP 119
Db 64 SVTTSIMYQYLATGKVDGTSSWVPTADKPYEKLTKFVDLGANYEGTIQGFVPSVYP 123
Qy 120 ISSISELKGKDKFKNMIGIDAGAGTQIVTEQALNYGLSKEYELVPSSESVMLASLDSS 179
Db 124 ISSISELKGKDKFKNMIGIDAGAGTQIVTEQALDYGLSKEYELVPSSESVMLASLDSS 183
Qy 180 SKRNEWILVPLMKPHWAFSRDYDKFLDDPDLIMGGIESVHTLVRLGLENDPDAYVVDH 239
Db 184 AKRNEWILVPLMKPHWAFSRDYDKFLDDPDLIMGGIESVHTLVRLGLENDPDAYVVDH 243
Qy 240 HFYWSDDLILPLMDKNDKPGKEYNRAVEFVEKNKEIVKTWVPEKYKTLFD 290
Db 244 HFYWNDDLILPLMDRNDKPGKEYNRAVEFVEKNKEIVKRWVPEKYKTLFD 294

RESULT 3
Q72903 DESVH
ID Q72903 DESVH PRELIMINARY; PRT; 285 AA.
AC Q72903;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Glycine/betaine/L-proline ABC transporter, periplasmic-binding
```

```
protein.
DE OrderedLocusNames=DVU2297;
GN Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
OS 8303).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBI_TaxID=882;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15077118; DOI=10.1038/nbt959;
RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Davidsen T.M., Zafar N., Zhou L., Radune D.,
RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Uitterback T.R.,
RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
RT Desulfovibrio vulgaris Hildenborough.";
RL Nat. Biotechnol. 22:554-559(2004).
DR EMBL; AE017316; AAS96770.1; -; Genomic_DNA.
DR TIGR; DVU2297; -.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007210; OpuAC_ABC.
DR Pfam; PF04069; OpuAC; 1.
KW Complete proteome.
SQ SEQUENCE 285 AA; 31240 MW; 7C18644EDDE6EE47 CRC64;

Query Match 33.4%; Score 510; DB 2; Length 285;
Best Local Similarity 37.0%; Pred. No. 1.1e-30;
Matches 105; Conservative 63; Mismatches 112; Indels 4; Gaps 4;

Qy 1 MYKFLFPIIFMF-LSCDEKSKSNLKSXKIGYVNWGGETAATNVLKVFFE-KMGYNAEI 58
Db 1 MKKIILLAVFLAFTGAQAQAKDPRIAYVEWDCARATSNLVKAAIEDRLHRKVEL 60
Qy 59 FSVTTSIMYQYLASGKIDGTSSWVPTADKPYEKLTKFVDLGANYEGTIQGFVPSV 118
Db 61 LPVSAGAMMAVASGDVDTATVMTLPTVTHGDLKRLGKVRDLGFLVDGARGLVAVPDV 120
Qy 119 PISSISELKGKDKFKNMIGIDAGAGTQIVTEQALNYGLSKEYELVPSSESVMLASLD 178
Db 121 TVNSIAELDAARDPFGRIIGIDPGAGLMRSEDAIKAYGL-KKQLVVGSGATWTAALA 179
Qy 179 SSIKRNEWILVPLMKPHWAFSRDYDKFLDDPDLIMGGIESVHTLVRLGLENDPDAYV 238
Db 180 DAIRKKEIVVTWSPHMFGRWQLKYLDDPKGALGGMESIHTVVRKGLDKDMPEVFAFL 239
Qy 239 DHFYWSDDLIL-PLMDKNDKPGKEYNRAVEFVEKNKEIVKTWV 281
Db 240 DRFAYADTADTQTLMAWNNEEGADPLTNARRFMKEHPALVDWSL 283

RESULT 4
Q8TNY0 METAC
ID Q8TNY0 METAC PRELIMINARY; PRT; 315 AA.
AC Q8TNY0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glycine betaine/L-proline ABC transporter, solute-binding
DE protein.
GN Name=proX; OrderedLocusNames=MA2147;
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238; DOI=10.1101/gr.223902;
```


RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,
RA Linton L., McKernan P., McKernan K., Talamas J., Titrill A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.T.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.,
RT "The genome of Methanohalobium acetivorans reveals extensive metabolic
RT and physiological diversity,"
RL Genome Res. 12:532-542(2002).
DR EMBL; AE010901; AAM05545.1; -; Genomic_DNA.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007210; OpuAC_ABC.
DR Pfam; PF04069; OpuAC; 1.
KW Complete proteome.
SQ SEQUENCE 315 AA; 34613 MW; 262494BFCT9649BA CRC64;

Query Match 33.1%; Score 505; DB 2; Length 315;
Best Local Similarity 41.6%; Pred. No. 3e-30;
Matches 112; Conservative 42; Mismatches 107; Indels 8; Gaps 5;

Qy 19 KSSNNKSVKIGYVWGGTAATNLVKVPEKGY-NAEIPSVTTSIMYOVLASGKIDG 77
Db 49 EASELSPTVIGYVWGGTAATNLVKVPEKGY-NAEIPSVTTSIMYOVLASGKIDG 108

Qy 78 TVSSVPTADKFFYEKLTKFVLDGANYEGTIGQFVPSYVPISSISLKGKGDKNKM 137
Db 109 TTSAWLPYTHQSWEAYGQDLSVQTNLEDCKGLVPSYVITDSIELNSEKDKFNQI 168

Qy 138 IGIDAGATQIVTEQALNYGLSKYELVPSSEVMSLASLSSIKRNEWILVPLKPHWA 197
Db 169 IGIDPGAGIMQASETAITDYDL--DMELVSGSSAAATTTSLKSIDSEEWTVVTLSPHWA 226

Qy 198 FSRVDIKFLDPPDLIMGGIESVHTLVRLGLENDDFDAYVDFHFWMS-DDLILPLMD-KN 255
Db 227 FNRDLKYLDDPKGAYGDADHVTGLRLGLEENKPNLGLTRFQWTHDDIQTVNMMDIEN 286

Qy 256 DKEPKGYRNVAFVEKKEIKVKTWVPEK 284
Db 287 GTAPEA---AANWVANNPEKNEWIGEK 312

RESULT 5
Q926L5 LISIN PRELIMINARY; PRT; 289 AA.
AC Q926L5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Pseudogene, similar to glycine-betaine binding (ABC transporter).
GN OrderedLocusNames=pli0045;
OS Listeria innocua.
OG Plasmid pI100.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Blocker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetoui F., Couve E., de Daruvar A., Deloux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fshi H., Garcia-Lopez N., Hain T., Hauf J., Jackson D.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunat F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,

RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomics of Listeria species,"
RL Science 294:849-852(2001).
DR EMBL; AL592102; CAC42043.1; -; Genomic_DNA.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007210; OpuAC_ABC.
DR Pfam; PF04069; OpuAC; 1.
KW Complete proteome; Plasmid.
SQ SEQUENCE 289 AA; 31737 MW; FASD1279BB86508F CRC64;

Query Match 32.4%; Score 494.5; DB 2; Length 289;
Best Local Similarity 37.9%; Pred. No. 1.7e-29;
Matches 106; Conservative 54; Mismatches 107; Indels 13; Gaps 5;

Qy 9 IIFNPLSCDEKSKNLKS-----VKIGYVWGGTAATNLVKVPEKGYNAEIPSVTTS 64
Db 15 LLFTFGSEDSYKTTVGADSGGQTVTLATVNWSESIASNTVLAQVLKSAEFGNVQITVDPA 74

Qy 65 IMYOVLASGKIDGTVSSVPTADKFFYEKLTKFVLDGANYEGTIGQFVPSYV-PISSI 123
Db 75 IMFSSVABEQGLDAMVGGVPTTHQAYAEKYGDSVVDLGANLEGAISALTVPYTMEDINSI 134

Qy 124 SELKGGKPKFKMIGIDAGAGTQIVTEQALNYGLSKYELVPSSEVMSLASLSSIKR 183
Db 135 TDLTDEND---STTATPEGAVTNSAQNAKEVDNLSDWEVSSTGAMTAEILDQALNN 191

Qy 184 NEWILVPLKPHWAFSRYDIKFLDPPDLIMGGIESVHTLVRLGLENDDFDAYVDFHFW 243
Db 192 EEDIVVVGKPHWFMFMDYDLKMLDPPENVFGYEEIHSYAREGLKEDNPAYKIIDNEFW 251

Qy 244 SDDLILPLMDK--NDKBPGEYRNVAFVEKKEIKVKTWV 281
Db 252 EVEDMSSVMESLATDVPEEAADN---WIEANREIVDGWL 288

RESULT 6
Q9A6X6 BACTN PRELIMINARY; PRT; 277 AA.
ID Q9A6X6 BACTN PRELIMINARY;
AC Q9A6X6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Glycine betaine-binding protein.
GN OrderedLocusNames=B71749;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=12663928; DOI=10.1126/science.1080029;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.,
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis,"
RL Science 299:2074-2076(2003).
DR EMBL; AE016933; AA076856.1; -; Genomic_DNA.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007210; OpuAC_ABC.
DR Pfam; PF04069; OpuAC; 1.
KW Complete proteome.
SQ SEQUENCE 277 AA; 31056 MW; B020C694CE12A21A CRC64;

Query Match 31.3%; Score 478.5; DB 2; Length 277;
Best Local Similarity 34.9%; Pred. No. 2.7e-28;
Matches 98; Conservative 59; Mismatches 119; Indels 5; Gaps 2;

RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20318987; PubMed=10860977; DOI=10.1073/pnas.97.13.7102;
RA van der Heide T., Poolman B.;
RT "Osmoregulated ABC-transport system of Lactococcus lactis senses water
stress via changes in the physical state of the membrane.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:7102-7106(2000).
DR EMBL; AF234619; AAF37879.1; -, Genomic_DNA.
DR PIR; C86806; C86806.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0005450; F:protein translocase activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0009306; P:protein secretion; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD transp.;
DR InterPro; IPR007210; OpuAC_ABC.
DR PRINTS; PR00303; SECYTRNLCASE.
DR Pfam; PF00528; BPD transp_1; 1.
DR PROSITE; PS50928; ABC_TM1; 1.
DR PROSITE; PS50928; ABC_TM1; 1.
SQ SEQUENCE 573 AA; 61952 MW; DFB59CEAE8462A0F CRC64;
Query Match 29.5%; Score 451; DB 2; Length 573;
Best Local Similarity 34.9%; Pred. No. 8.3e-26;
Matches 99; Conservative 64; Mismatches 99; Indels 22; Gaps 7;
QY 6 LFFIIFMFLSCDEKSSKNLKVIGYVNWGGETAATNVLKVPFKMGYNAEIFSVTTSI 65
DB 302 LALIIGAFSGMSFGKTASD--KKVDLVYVNWDESEASINVLTAQMKHGFVKTALDNAV 360
QY 66 MYQYLASGKIDGTSSWVPADKFFYEKLTKFVD--LGANYEGTIGQFVVPSPYVPISSIS 124
DB 361 AWQTVANGQADGWSAWLPNTHKTQWKY--GKSVDLGLGNLKGAKGVFPVPSYMNVSIE 419
QY 125 ELKGGDKFKPKMIGIDAGAGTQIVTEQALNYGLSKEYELVPSSSEVMLASLDSISIKRN 184
DB 420 DLTNQANK---TITGIEPGAGVMAASEKTLSYDNLKDWKLVPPSSSGAMTVALGEAIKQH 476
QY 185 EWILVPLKPHWAFSRDYIKFLDDPDLMGGIESVHTLVRLGLENDDFDYYVDFHFW- 243
DB 477 KDIVITGSPHWMFNKYDLKADPKGTMTGTSININTIVRKLKXENPEAYKVLDFKNWT 536
QY 244 SDDLILPLMD--KNDKEPGKEYRNAVEFVEKKEIKVKTWVPEYK 286
DB 537 TKDMEAVMLDIQNGKTP-----EEAAKNWIKDHQK 566
RESULT 12
Q7DAU8 LACLA PRELIMINARY; PRT; 573 AA.
ID Q7DAU8 LACLA PRELIMINARY; PRT; 297 AA.
AC Q7DAU8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Betaine ABC transporter permease and substrate binding protein.
GN Name=busAB; OrderedLocNames=IL1451;
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471; DOI=10.1101/gr.CR-1697R;
RA Bolotin A., Winkler P., Mager S., Jallion O., Malarie K.,
RT Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753(2001).
DR EMBL; AE006375; AA05549.1; -, Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0015450; F:protein translocase activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0009306; P:protein secretion; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD transp.;
DR InterPro; IPR007210; OpuAC_ABC.
DR PRINTS; PR00303; SECYTRNLCASE.
DR Pfam; PF00528; BPD transp_1; 1.
DR PROSITE; PS50928; ABC_TM1; 1.
DR PROSITE; PS50928; ABC_TM1; 1.
KW Complete proteome.
SQ SEQUENCE 573 AA; 61952 MW; DFB59CEAE8462A0F CRC64;
Query Match 29.5%; Score 451; DB 2; Length 573;
Best Local Similarity 34.9%; Pred. No. 8.3e-26;
Matches 99; Conservative 64; Mismatches 99; Indels 22; Gaps 7;
QY 6 LFFIIFMFLSCDEKSSKNLKVIGYVNWGGETAATNVLKVPFKMGYNAEIFSVTTSI 65
DB 302 LALIIGAFSGMSFGKTASD--KKVDLVYVNWDESEASINVLTAQMKHGFVKTALDNAV 360
QY 66 MYQYLASGKIDGTSSWVPADKFFYEKLTKFVD--LGANYEGTIGQFVVPSPYVPISSIS 124
DB 361 AWQTVANGQADGWSAWLPNTHKTQWKY--GKSVDLGLGNLKGAKGVFPVPSYMNVSIE 419
QY 125 ELKGGDKFKPKMIGIDAGAGTQIVTEQALNYGLSKEYELVPSSSEVMLASLDSISIKRN 184
DB 420 DLTNQANK---TITGIEPGAGVMAASEKTLSYDNLKDWKLVPPSSSGAMTVALGEAIKQH 476
QY 185 EWILVPLKPHWAFSRDYIKFLDDPDLMGGIESVHTLVRLGLENDDFDYYVDFHFW- 243
DB 477 KDIVITGSPHWMFNKYDLKADPKGTMTGTSININTIVRKLKXENPEAYKVLDFKNWT 536
QY 244 SDDLILPLMD--KNDKEPGKEYRNAVEFVEKKEIKVKTWVPEYK 286
DB 537 TKDMEAVMLDIQNGKTP-----EEAAKNWIKDHQK 566
RESULT 13
Q8EKK8 OCEIH PRELIMINARY; PRT; 297 AA.
ID Q8EKK8 OCEIH PRELIMINARY; PRT; 297 AA.
AC Q8EKK8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glycine betaine ABC transporter glycine betaine-binding protein.
GN OrderedLocNames=OB3454;
OS Oceanobacillus thelyensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
OX NCBI_TaxID=182710;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HT831 / DSM 14371 / JCM 11309;
RX MEDLINE=2220767; PubMed=1223376; DOI=10.1093/nar/gkf526;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus thelyensis isolated from the Iheya
Ridge and its unexpected adaptive capabilities to extreme
environments.";
RL Nucleic Acids Res. 30:3927-3935(2002).
DR EMBL; BA000028; BAC15410.1; -, Genomic_DNA.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0030001; P:metal ion transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR006121; HeavyMe transp.
DR InterPro; IPR007210; OpuAC_ABC.
DR Pfam; PF04069; OpuAC; 1.
KW Complete proteome.
SQ SEQUENCE 297 AA; 32068 MW; C01B0FAC9BDCBE1A CRC64;

```

Query Match      29.3%; Score 447; DB 2; Length 297;
Best Local Similarity 34.1%; Pred. No. 7.5e-26;
Matches 92; Conservative 57; Mismatches 103; Indels 12; Gaps 3;

QY 18 EKKSKNLKSVKIGVYNNWGGETAATNLVKVPEKMGYNABIFSVTTSIMYQYLASGKIDG 77
DB 34 EGDSEAKKEIEIAYVENDTEVASTHVIGKVLDELGYVELTPLDNLAIWEAVANGAEGAD 93
QY 78 TVSSWVPADKPYEKLTKFVDLGANYEGTTQGFVVPYSYV-PISSISELKGKDKFKNM 137
DB 94 MVAALPNTARQYESYGDQVESLGNLEGAKIGLVPEYMDVNSIEDL---SDEAGQTI 150
QY 138 IGIDAGAGTQIVTEQALNYGLSKEYELVPSSEVNLASLDSSIKENWILVPLKPHWA 197
DB 151 TGIEPGVAAASEDAVETYNLQWSVQTSSEGAATLGSADNEBEEIIVTGSHPWK 210
QY 198 FSRYDIKELDDPLIMGGIESVHTLVRLGLENDPDAVYVDFHFWYS-----DDLPLMD 253
DB 211 FQAYDLKYLEDEPGVFGDAETIETWVRGLEEDMPEATVLDNFQDFTAAMEEVMNLISE 270
QY 254 KNDKEPGKEYRNAVEFVEKNEIVKTVWPE 283
DB 271 GADPE-----QAADWVEANQDKVDWTKD 295

RESULT 14
Q8K8R4_STRP3 PRELIMINARY; PRT; 561 AA.
AC Q8K8R4;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DE Putative glycine-betaine binding permease protein.
GN Name=opuABC; OrderedLocNames=SpvM3_0144;
OS Streptococcus pyogenes (serotype M3)
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301448;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206; DOI=10.1073/pnas.152298499;
RA Beres S.B., Sylva G.B., Barbian K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Perkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus: phage-encoded toxins, the high-virulence phenotype, and clone emergence.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).
DR EMBL; AB014139; AAM78751.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0015450; F:protein translocase activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0009306; P:protein secretion; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD transp.
DR InterPro; IPR007210; OpuAC ABC.
DR Pfam; PF04069; OpuAC; 1.
DR PRINTS; PR00303; SECYTNLCASE.
DR PROSITE; PS50928; ABC_TM1; 1.
KW Complete proteome.
SQ SEQUENCE 561 AA; 61568 MW; 59BAE2C139AA55C2 CRC64;

Query Match      29.0%; Score 442.5; DB 2; Length 561;
Best Local Similarity 35.8%; Pred. No. 3.6e-25;
Matches 92; Conservative 55; Mismatches 103; Indels 7; Gaps 5;

QY 26 KSVKIGYNNWGGETAATNLVKVPEKMGYNABIFSVTTSIMYQYLASGKIDGTSSWVPT 85
DB 322 ETVNIAYVQWSEVASTHVIAEVLKNEGYHVTLPDNLAVNMQTVANGNADFSTSAWLPV 381
QY 86 ADKFFYEKLTKFVDLGANYEGTTQGFVVPYSYV-PISSISELKGKDKFKNMIGIDAGA 144

Query Match      29.0%; Score 442.5; DB 2; Length 575;
Best Local Similarity 35.8%; Pred. No. 3.7e-25;
Matches 92; Conservative 55; Mismatches 103; Indels 7; Gaps 5;

QY 26 KSVKIGYNNWGGETAATNLVKVPEKMGYNABIFSVTTSIMYQYLASGKIDGTSSWVPT 85
DB 322 ETVNIAYVQWSEVASTHVIAEVLKNEGYHVTLPDNLAVNMQTVANGNADFSTSAWLPV 381
QY 86 ADKFFYEKLTKFVDLGANYEGTTQGFVVPYSYV-PISSISELKGKDKFKNMIGIDAGA 144

```

Db	382	THGQQYQYKSKLDDLGNLKGTLGLAVPKYMTDVNSIEDLSKQADQ---KITGIEPGA	438
QY	145	GTQIVTEQALNTYGLSKEYELVPSSSEVWLASLDSIKNNEWILVPLWKPHWAFSRDYDK	204
Db	439	GIMAAAKKTLKEYHNLSSWELVAASTGAMTTSLDQAIIKKKDPVVVTAMSPHMMFAKYDLK	498
QY	205	FLDDPDLIMGGIESVHTLVRLGLENDPDAYVYVDFHYWSD-DLJLPLMDKNDKEPGKEY	263
Db	499	YLDKPKETFGSTENINTIARKGLKKDLNVYKIIDKFHWTQKDMAAVMLDIN-KGMSPE-	556
QY	264	RNAVEFEKNEKIVKTW	280
Db	557	AAAKKWEANKSVSSW	573

Search completed: January 24, 2006, 19:56:44
Job time : 63.3148 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:21:03 ; Search time 45.5886 Seconds
(without alignments)
1966.133 Million cell updates/sec

Title: US-10-688-058-10
Perfect score: 1090
Sequence: 1 RKFPFDLNDVTVMPLQK.....QFINYCSTLSKFFVDDNIM 204

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: geneseq1980s.*
2: geneseq1990s.*
3: geneseq2000s.*
4: geneseq2001s.*
5: geneseq2002s.*
6: geneseq2003as.*
7: geneseq2003bs.*
8: geneseq2004s.*
9: geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1090	100.0	204	ADU98751	Adu98751 Borrelia
2	1090	100.0	377	ADU98753	Adu98753 Borrelia
3	152.5	14.0	385	ABM70978	Abm70978 Staphyloc
4	151.5	13.9	385	ABJ18960	Abj18960 Pathogen
5	150	13.8	375	ABE40038	Aeb40038 L. pneumo
6	150	13.8	382	ABE36646	Aeb36646 L. pneumo
7	145	13.3	156	AY20035	Aay20035 B. burgdo
8	145	13.3	176	AY20034	Aay20034 B. burgdo
9	141.5	13.0	408	ABP27526	Abp27526 Streptoco
10	137.5	12.6	379	ABB54455	Abb54455 Lactococc
11	137	12.6	376	ADV87963	Adv87963 Streptoco
12	137	12.6	376	ADV79216	Adv79216 Streptoco
13	137	12.6	376	ADV81406	Adv81406 Streptoco
14	136	12.5	376	ABP27525	Abp27525 Streptoco
15	135	12.4	381	AEA49154	Aea49154 L. rhanno
16	134	12.3	376	ABU01854	Abu01854 S. pneumo
17	134	12.3	376	ADK46982	Adk46982 Streptoco
18	134	12.3	409	ADR95677	Adr95677 Novel S.
19	134	12.3	409	AEA59547	Aea59547 Streptoco
20	118.5	10.9	295	AY37169	Aay37169 Protein i
21	114.5	10.5	385	ABB48803	Abb48803 Listeria
22	112	10.3	391	ABP38845	Abp38845 Staphyloc
23	111	10.2	393	ADC97174	Adc97174 E. faeciu
24	107.5	9.9	442	AAB93553	Aab93553 Human pro

25	107.5	9.9	442	8	ABM81907	Abm81907 Tumour-as
26	105.5	9.7	401	9	ABM95708	Abm95708 M. xanthu
27	105	9.6	424	7	ABO71949	AbO71949 Pseudomon
28	104.5	9.6	397	7	ADH88000	Adh88000 Enterococ
29	102.5	9.4	394	2	AY34978	Aay34978 Chlamydia
30	100.5	9.2	377	6	ABM67734	Abm67734 Photorhab
31	99.5	9.1	409	4	AAU67750	Aau67750 Propionib
32	99.5	9.1	409	6	ABM64269	Abm64269 Propionib
33	99.5	9.1	429	6	ABM65004	Abm65004 Klebsiell
34	98	9.0	387	7	ABO65431	AbO65431 Klebsiell
35	97	8.9	548	8	ADS08193	AdS08193 Staphyloc
36	96.5	8.9	382	7	ADF04923	Adf04923 Bacterial
37	95.5	8.8	442	4	AB95029	Ab95029 Human pro
38	93	8.5	350	4	ABB62280	Abb62280 Drosophil
39	93	8.5	4987	2	AAR10834	Aar10834 Rianodin
40	91.5	8.4	605	6	ABU26712	Abu26712 Protein e
41	91	8.3	1103	6	ABU48942	Abu48942 Enterohae
42	90	8.3	605	7	ADC00089	Adc00089 Plasmodiu
43	89	8.2	2024	8	ADP25444	Adp25444 B. burgdo
44	88.5	8.1	288	2	AAI19835	Aay19835 B. burgdo
45	88.5	8.1	310	2	AAI19834	Aay19834 B. burgdo

ALIGNMENTS

RESULT 1
ADU98751
ID ADU98751 standard; protein; 204 AA.
AC ADU98751;
XX
XX
DT 24-FEB-2005 (first entry)
XX
DE Borrelia burgdorferi antigenic polypeptide seqid 10.
XX
KW antibacterial; vaccine; immune stimulation; immunity; antigen;
KW DNA library.
XX
OS Borrelia burgdorferi.
XX
XX
PN WO2004103269-A2.
XX
PD 02-DEC-2004.
XX
PF 17-OCT-2003; 2003WO-US033056.
XX
PR 18-OCT-2002; 2002US-0419401P.
XX
PA (MACR-) MACROGENICS INC.
(TEXA) UNIV TEXAS SYSTEM.
XX
PI Sykes KF, Hale KS, Johnston SA;
XX
DR WPI; 2004-834155/82.
XX
DR N-PSDB; ADU98750.
XX
PT Immunizing a subject against Borrelia burgdorferi infection comprises
providing to the subject at least one Borrelia antigen or its fragment.
XX
PS Claim 27; SEQ ID NO 10; 121pp; English.
XX
CC The invention describes a method of immunizing a subject comprising
providing to the subject at least one Borrelia antigen or its fragment to
induce an immune response. Also described are: an isolated polynucleotide
comprising a sequence having at least 17 contiguous nucleotides in common
with a sequence not given in the specification; an isolated polypeptide
having at least 5 consecutive amino acids of the sequence not given in
the specification; a vaccine composition comprising at least one Borrelia
antigen or at least one polynucleotide encoding a Borrelia antigen;
screening for at least one test polypeptide or test polynucleotide
encoding a polypeptide for an ability to produce an immune response;
preparing a vaccine; vaccinating a subject; treating a subject infected

CC with a pathogen; passive immunization; vaccination; detecting Herpesvirus
CC ; and detecting antibodies to a herpesvirus. The method is useful in
CC immunizing a subject against Borrelia burgdorferi infection. This is the
CC amino acids sequence of a Borrelia burgdorferi polypeptide that may be
CC useful in immunizing a subject against infection by a member of the
CC Borrelia species.

XX Sequence 204 AA;

SQ Query Match 100.0%; Score 1090; DB 8; Length 204;
Best Local Similarity 100.0%; Pred. No. 1e-106;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKPPFDLNDMTVMPLQKSHLKRDLKELLSYMPHEICFSDFCIEEGFVLRDFDINSID 60

Db 1 RKPPFDLNDMTVMPLQKSHLKRDLKELLSYMPHEICFSDFCIEEGFVLRDFDINSID 60

Qy 61 SEKLWFCALCECLSENGYINYEITNFALKGHESRHNKLNWELKPHLGLGLYAVSLFCNDK 120

Db 61 SEKLWFCALCECLSENGYINYEITNFALKGHESRHNKLNWELKPHLGLGLYAVSLFCNDK 120

Qy 121 NNNVRALIRKTGSFVKANNHLVTPELLEDEFFVYHFIQGLGTIGQVSLRALRLRFYNE 180

Db 121 NNNVRALIRKTGSFVKANNHLVTPELLEDEFFVYHFIQGLGTIGQVSLRALRLRFYNE 180

Qy 181 KQFFQFINYCSTLSKKFVFDNDIM 204

Db 181 KQFFQFINYCSTLSKKFVFDNDIM 204

RESULT 2

ADU98753
ID ADU98753 standard; protein; 377 AA.

XX AC ADU98753;

XX AC 24-FEB-2005 (first entry)

XX Borrelia burgdorferi antigenic polypeptide seqid 12.

XX antibacterial; vaccine; immune stimulation; immunity; antigen;
XX DNA library.

XX Borrelia burgdorferi.

XX WO2004103269-A2.

XX 02-DEC-2004.

XX 17-OCT-2003; 2003WO-US033056.

XX 18-OCT-2002; 2002US-0419401P.

XX (MACR-) MACROGENICS INC.

XX (TEXA) UNIV TEXAS SYSTEM.

XX Sykes KF, Hale KS, Johnston SA;

XX WPI; 2004-834155/82.

XX N-PSDB; ADU98752.

XX Immunizing a subject against Borrelia burgdorferi infection comprises
XX providing to the subject at least one Borrelia antigen or its fragment.

XX Claim 27; SEQ ID NO 12; 121pp; English.

XX The invention describes a method of immunizing a subject comprising
XX providing to the subject at least one Borrelia antigen or its fragment to
XX induce an immune response. Also described are: an isolated polynucleotide
XX comprising a sequence having at least 17 contiguous nucleotides in common
XX with a sequence not given in the specification; an isolated polypeptide
XX having at least 5 consecutive amino acids of the sequence not given in
XX the specification; a vaccine composition comprising at least one Borrelia

CC antigen or at least one polynucleotide encoding a Borrelia antigen;
CC screening for at least one test polypeptide or test polynucleotide
CC encoding a polypeptide for an ability to produce an immune response;
CC preparing a vaccine; vaccinating a subject; treating a subject infected
CC with a pathogen; passive immunization; vaccination; detecting Herpesvirus
CC ; and detecting antibodies to a herpesvirus. The method is useful in
CC immunizing a subject against Borrelia burgdorferi infection. This is the
CC amino acids sequence of a Borrelia burgdorferi polypeptide that may be
CC useful in immunizing a subject against infection by a member of the
CC Borrelia species.

XX Sequence 377 AA;

SQ Query Match 100.0%; Score 1090; DB 8; Length 377;

Best Local Similarity 100.0%; Pred. No. 2.3e-106;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKPPFDLNDMTVMPLQKSHLKRDLKELLSYMPHEICFSDFCIEEGFVLRDFDINSID 60

Db 141 RKPPFDLNDMTVMPLQKSHLKRDLKELLSYMPHEICFSDFCIEEGFVLRDFDINSID 200

Qy 61 SEKLWFCALCECLSENGYINYEITNFALKGHESRHNKLNWELKPHLGLGLYAVSLFCNDK 120

Db 201 SEKLWFCALCECLSENGYINYEITNFALKGHESRHNKLNWELKPHLGLGLYAVSLFCNDK 260

Qy 121 NNNVRALIRKTGSFVKANNHLVTPELLEDEFFVYHFIQGLGTIGQVSLRALRLRFYNE 180

Db 261 NNNVRALIRKTGSFVKANNHLVTPELLEDEFFVYHFIQGLGTIGQVSLRALRLRFYNE 320

Qy 181 KQFFQFINYCSTLSKKFVFDNDIM 204

Db 321 KQFFQFINYCSTLSKKFVFDNDIM 344

RESULT 3

ABM70978

ID ABM70978 standard; protein; 385 AA.

XX AC ABM70978;

XX XX 20-NOV-2003 (first entry)

XX Staphylococcus aureus protein #218.

XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
XX enzymatic assay; antibiotic target.

XX Staphylococcus aureus.

XX WO200294868-A2.

XX 28-NOV-2002.

XX 27-MAR-2002; 2002WO-IB002637.

XX 27-MAR-2001; 2001GB-00007661.

XX (CHIR-) CHIRON SPA.

XX Masignani V, Mora M, Scarselli M;

XX WPI; 2003-120786/11.

XX N-PSDB; ACF72538.

XX New Staphylococcus aureus protein, useful as a vaccine for treating or
XX preventing Staphylococcal infection, specifically an infection caused by
XX S. aureus, e.g. sepsis.

XX Claim 1; SEQ ID NO 436; 49pp; English.

XX The invention relates to novel genes and encoded proteins from
XX Staphylococcus aureus. A composition comprising the S. aureus protein, a
XX nucleic acid encoding the protein, or an antibody to the protein, is


```

XX New genome of Legionella pneumophila Paris strain and derived
PT polypeptides, useful for detection or identification of the strain and
PT for treatment and prevention of infections.
XX
XX Claim 3; SEQ ID NO 4370; 660pp; English.
XX
XX The invention relates to an isolated or purified nucleotide sequences (I)
CC from Legionella pneumophila Paris strain. (II), and their related
CC sequences or fragments, are useful as primers and probes for detection
CC and amplification, including differentiation between the Paris and
CC Philadelphia strains of Legionella pneumophila and to prepare recombinant
CC (hybrid) polypeptides (II). (II) are also useful for preparation of
CC specific antibodies (Ab), also used for detection/identification of
CC Legionella, and some (I), specifically those involved in synthesis of
CC surface proteins, are targets for identification of inhibitors. (II), or
CC vectors that contain (I), are useful as vaccines and immunogenic
CC compositions, for treatment and prevention of infections by L.
CC pneumophila. The present sequence represents the amino acid sequence of a
CC L. pneumophila protein.
XX
XX Sequence 375 AA;
SQ
Query Match 13.8%; Score 150; DB 9; Length 375;
Best Local Similarity 28.3%; Pred. No. 6.9e-07;
Matches 43; Conservative 21; Mismatches 40; Indels 48; Gaps 4;
Qy 1 RKFPD-LNIDMTVMPLQKSHLKRDLKELLSYMPHICF----- 40
Db 150 RKAGFDNLNLDIMHSLPNQSVTQGLQDLKTALSYQPEHLSWYQLTIEPNTVFKHTPLP 209
Qy 41 ---SDFICEEGFVLDPDINSIDSEKLFWFCALCLESNGYINYEITNFALKGHESRNKL 97
Db 210 SEEDYLLEEGFAL-----LHNSGNYRYEISAFSPKPKQARHNIN 250
Qy 98 NWELKPHLGLGLYAVSLFLFCNDKNNVRLAIR 129
Db 251 YWLFQDYLIGIGAGA-----HGKMTTPNAIR 276
RESULT 6
AEB36646
ID AEB36646 standard; protein; 382 AA.
XX
AC AEB36646;
XX
08-SEP-2005 (first entry)
XX
L. pneumophila protein SEQ ID NO 978.
XX
detection; infection; Antibacterial; Vaccine.
XX
Legionella pneumophila.
XX
WO2005049642-A2.
XX
02-JUN-2005.
XX
23-SEP-2004; 2004WO-IB003578.
XX
21-NOV-2003; 2003FR-00013687.
XX
(INSP ) INST PASTEUR.
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA (UTLY-) UNIV LYON 1 BERNARD CLAUDE.
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
Buchrieser C, Tichit M, Etienne J, Ma L, Cazalet C, Glaser P;
PI Rusniok C, Bouchier C, Zidane N, Magnier A, Kunst F, Vandenesch F;
PI Jarraud S;
XX
WPI; 2005-388305/40.
XX

```

```

PT New genome of Legionella pneumophila Paris strain and derived
PT polypeptides, useful for detection or identification of the strain and
PT for treatment and prevention of infections.
XX
XX Claim 3; SEQ ID NO 978; 660pp; English.
XX
XX The invention relates to an isolated or purified nucleotide sequences (I)
CC from Legionella pneumophila Paris strain. (II), and their related
CC sequences or fragments, are useful as primers and probes for detection
CC and amplification, including differentiation between the Paris and
CC Philadelphia strains of Legionella pneumophila and to prepare recombinant
CC (hybrid) polypeptides (II). (II) are also useful for preparation of
CC specific antibodies (Ab), also used for detection/identification of
CC Legionella, and some (I), specifically those involved in synthesis of
CC surface proteins, are targets for identification of inhibitors. (II), or
CC vectors that contain (I), are useful as vaccines and immunogenic
CC compositions, for treatment and prevention of infections by L.
CC pneumophila. The present sequence represents the amino acid sequence of a
CC L. pneumophila protein.
XX
XX Sequence 382 AA;
SQ
Query Match 13.8%; Score 150; DB 9; Length 382;
Best Local Similarity 28.3%; Pred. No. 7.1e-07;
Matches 43; Conservative 21; Mismatches 40; Indels 48; Gaps 4;
Qy 1 RKFPD-LNIDMTVMPLQKSHLKRDLKELLSYMPHICF----- 40
Db 157 RKAGFDNLNLDIMHSLPNQSVTQGLQDLKTALSYQPEHLSWYQLTIEPNTVFKHTPLP 216
Qy 41 ---SDFICEEGFVLDPDINSIDSEKLFWFCALCLESNGYINYEITNFALKGHESRNKL 97
Db 217 SEEDYLLEEGFAL-----LHNSGNYRYEISAFSPKPKQARHNIN 257
Qy 98 NWELKPHLGLGLYAVSLFLFCNDKNNVRLAIR 129
Db 258 YWLFQDYLIGIGAGA-----HGKMTTPNAIR 283
RESULT 7
AAAY20035
ID AAAY20035 standard; protein; 156 AA.
XX
AC AAAY20035;
XX
19-JUL-1999 (first entry)
XX
B. burgdorferi antigenic protein, t229.aa.
XX
Antigenic protein; vaccine; Lyme disease; infection; detection.
XX
Borrelia burgdorferi.
XX
WO9859071-A1.
XX
30-DEC-1998.
XX
18-JUN-1998; 98WO-US012718.
XX
20-JUN-1997; 97US-0050359P.
PR 22-JUL-1997; 97US-0053344P.
PR 22-JUL-1997; 97US-0053377P.
PR 03-SEP-1997; 97US-0057483P.
XX
(HUMA-) HUMAN GENOME SCI INC.
PA (MEDI-) MEDIMUNE INC.
XX
Choi GH, Erwin AL, Hanson MS, Lathigra R;
PI WPI; 1999-189980/16.
DR N-ESDB; AAX61732.
XX
New isolated Borrelia burgdorferi nucleic acids - used to develop
PT

```

PT products for the diagnosis, prevention and treatment of diseases caused
 XX by Borrelia, particularly Lyme disease.

PS Claim 12; Page 169; 275pp; English.

XX This sequence represents a Borrelia burgdorferi (Bb) protein of the
 CC invention, which is suitable for use in a vaccine. The Bb polypeptides
 CC can be used in vaccines for eliciting protective antibodies to members of
 CC the Borrelia genus, particularly for the use against Lyme disease in
 CC humans and animals. They can be used for preventing or attenuating an
 CC infection caused by a member of the Borrelia genus. The products can also
 CC be used for detection of members of the Borrelia genus

XX SQ Sequence 156 AA;

Query Match 13.3%; Score 145; DB 2; Length 156;
 Best Local Similarity 90.0%; Pred. No. 7.3e-07;
 Matches 27; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RKFPFDLNDMTVMPLQKKSHLKRDLKEL 30
 Db 121 RKFPFDLNDMTVMPLQKKSHLKRDLQRI 150

RESULT 8
 RAY20034
 ID AAY20034 standard; protein; 176 AA.

XX AC AAY20034;
 XX 19-JUL-1999 (first entry)

XX B. burgdorferi antigenic protein, f229.aa.
 DE Antigenic protein; vaccine; Lyme disease; infection; detection.
 XX Borrelia burgdorferi.

XX W09859071-A1.
 XX 30-DEC-1998.
 XX 18-JUN-1998; 98WO-US012718.

XX 20-JUN-1997; 97US-0050359P.
 XX 22-JUL-1997; 97US-0053344P.
 XX 22-JUL-1997; 97US-0053377P.
 XX 03-SEP-1997; 97US-0057483P.

XX (HUMA-) HUMAN GENOME SCI INC.
 FA (MEDI-) MEDIMMUNE INC.

XX Choi GH, Erwin AL, Hanson MS, Lathigra R;
 XX WPI; 1999-189980/16.
 DR N-PSDB; AAX61731.

XX New isolated Borrelia burgdorferi nucleic acids - used to develop
 PT products for the diagnosis, prevention and treatment of diseases caused
 PT by Borrelia, particularly Lyme disease.

XX Claim 12; Page 169; 275pp; English.

XX This sequence represents a Borrelia burgdorferi (Bb) protein of the
 CC invention, which is suitable for use in a vaccine. The Bb polypeptides
 CC can be used in vaccines for eliciting protective antibodies to members of
 CC the Borrelia genus, particularly for the use against Lyme disease in
 CC humans and animals. They can be used for preventing or attenuating an
 CC infection caused by a member of the Borrelia genus. The products can also
 CC be used for detection of members of the Borrelia genus

XX SQ Sequence 176 AA;

Query Match 13.3%; Score 145; DB 2; Length 176;
 Best Local Similarity 90.0%; Pred. No. 8.6e-07;
 Matches 27; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RKFPFDLNDMTVMPLQKKSHLKRDLKEL 30
 Db 141 RKFPFDLNDMTVMPLQKKSHLKRDLQRI 170

RESULT 9
 ABP27526
 ID ABP27526 standard; protein; 408 AA.

XX AC ABP27526;
 XX 02-JUL-2002 (first entry)

XX Streptococcus polypeptide SEQ ID NO 4228.

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 XX group A streptococcus; Streptococcus pyogenes; antibacterial;
 XX antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX Streptococcus pyogenes.
 XX W0200234771-A2.
 XX 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB004789.
 XX 27-OCT-2000; 2000GB-00026333.
 XX 24-NOV-2000; 2000GB-00028727.
 XX 07-MAR-2001; 2001GB-00005640.

XX (CHIR-) CHIRON SPA.
 XX (GENO-) INST GENOMIC RES.

XX Telford J, Maignani V, Margarit Y Rosi, Grandi G, Fraser C;
 XX Tettelin H;
 XX WPI; 2002-352536/38.
 XX N-PSDB; ABN68157.

XX New Streptococcus protein for the treatment or prevention of infection or
 PT disease caused by Streptococcus bacteria, such as meningitis, and for
 PT detecting a compound that binds to the protein.

XX Claim 1; Page 3580; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), AN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins

XX Sequence 408 AA;

Query Match 13.0%; Score 141.5; DB 5; Length 408;
 Best Local Similarity 25.1%; Pred. No. 6.1e-06;
 Matches 45; Conservative 32; Mismatches 91; Indels 11; Gaps 3;

Matches	47;	Conservative	36;	Mismatches	90;	Indels	24;	Gaps	6
Qy	6	DLMIDMTVNNPLOKXKHLKDLKELLSYMEPHICFSDFIGEEBGFVLDRF-----DN	57						
		: : :	:	: : :	:	:	:		
Db	153	NISIDLIALPGQTMDVRNSVAKALSLNTPHLSLYSLILEHHTVFMNKMRGKGLHPTE	212						
Qy	58	SIDSSEKLWFCALCESLGESNYINFTNFALKGHESHINKLNWELKPHLGLGLVAVSLLFC	117						
		: : :	:	: : :	:	:	:	:	
Db	213	DLEAEMFEYIISE--MERNGFHEYIEISNFTKPGFGRSHNLMYWDNVVEYGVGAGASGYL--	269						
Qy	118	NDKNKNNVALIRKTGP-----VKANNHLVTFELLEDELFVVYHFIOGLGTIOGVSLRA	171						
		: : :	:	: : :	:	:	:	:	
Db	270	----DGR--YRNRGPIQHVLKGVSIGNARLSSEVLSSKNMMEELFLGURKEGVSIGK	323						
Qy	172	LRLRFEYN-EKQFFQFI	187						
		: : :	:	: : :	:	:	:	:	
Db	324	FEQKFGTSFEKRYGIV	340						
RESULT	13								
ID	ADV81406								
XX	ADV81406	standard; protein; 376 AA.							
AC	ADV81406;								
DT	24-FEB-2005	(first entry)							
DE	Streptococcus agalactiae	protein, SEQ ID 2547.							
KW	Antibacterial; vaccine; bacterial infection.								
OS	Streptococcus agalactiae.								
FN	WO200292818-A2.								
FD	21-NOV-2002.								
PF	26-APR-2002; 2002WO-IB003059.								
PR	26-APR-2001; 2001FR-00005642.								
FA	(INSP) INST PASTEUR.								
PA	(CNRS) CNRS CENT NAT RECH SCI.								
PI	Glaser P, Rusniok C, Chevalier F, Frangeul L, Lalioui L;								
EI	Zouine M, Couve E, Buchrieser C, Poyart C, Trieu-Cuot P, Kunst F;								
XX	WPT; 2004-101891/11.								
DR									
XX									
FS	Genomic nucleotide sequences encoding polypeptides of Streptococcus								
PT	agalactiae for the development of vaccines, diagnostic tools, DNA chips								
PT	and identification of therapeutic targets.								
XX									
FS	Claim 6; SEQ ID NO 2547; 439pp; French.								
XX									
CC	The present invention relates to novel Streptococcus agalactiae								
CC	nucleotide sequences (I; ADV78860-ADV78998 and ADV83341-ADV85476) and								
CC	novel polypeptides (II; ADV78999-ADV81203 and ADV81205-ADV83340). The								
CC	nucleotide sequences encode polypeptides of S. agalactiae involved in the								
CC	synthesis of amino acids, cell membranes, intermediate (central								
CC	metabolism, energetic metabolism, fatty acid and phospholipid metabolism,								
CC	nucleotide metabolism including purines, pyrimidines and/or nucleosides,								
CC	regulatory functions, replication, transcription, translation, protein								
CC	transport, adaptation to atypical conditions, sensitivity to medicines								
CC	and/or analogues, functions related to transposons, biosynthesis of								
CC	cofactors, prosthetic groups and transporters, cell membrane proteins and								
CC	cellular machinery. (I) are useful for the detection and/or amplification								
CC	of nucleic acids. Pharmaceutical composition comprising (I) or (II) are								
CC	useful for treatment of a bacterial S. agalactiae infection. The complete								
CC	genome of Streptococcus agalactiae is given in ADV81204. Note: The								
CC	present patent is an equivalent for the basic patent FR2824074A1, which								
XX	contains only 2344 sequences.								
XX									

CC present patent is an equivalent for the basic patent FR28240/4A1, which
CC contains only 2344 sequences.
XX

```
SQ      Sequence 376 AA;
Query Match      12.6%; Score 137; DB 8; Length 376;
Best Local Similarity 23.9%; Pred. No. 1.7e-05;
Matches 47; Conservative 36; Mismatches 90; Indels 24; Gaps 6;

QY      6 DLNIDMTVMPLQKSHLRDLKELLSYMPHEICFSDFIGEEGFLVDRP-----DN 57
DB      153 NISDILYALPGQTMDDVRNVAKALSINPHLSLSYLILEHHTVFMNKRGRKHLHLPTE 212

QY      58 SIDSEKLMWFCALCELSNGVINYETNFALKGHESRHNKLNWELKPHLGLGLYAVSLFLC 117
DB      213 DLEAEMFEYIIE-MERNGFEHYEISNFTKPGFESRHNLMYWDNVEYGVGAGASGYL-- 269

QY      118 NDKNNVRALIRKTGSF-----VKANNHLVTPELLEDEFFVYHFIQGLGTIQGVSLRA 171
DB      270 ----DGIR--YRNRGPIQHYLKGVSNGNARLSEVLKSNEMEEELFLGLRKKEGVSIK 323

QY      172 LRLRFEYN-EKQFFQFI 187
DB      324 FEQKFGTSFEKRYGQIV 340

RESULT 14
ABP27525
ID      ABP27525 standard; protein; 376 AA.
XX
AC      ABP27525;
XX
XX
DT      02-JUL-2002 (first entry)
XX
DE      Streptococcus polypeptide SEQ ID NO 4226.
XX
KW      Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW      group A streptococcus; Streptococcus pyogenes; antibacterial;
KW      antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS      Streptococcus agalactiae.
XX
FN      WO200234771-A2.
XX
PD      02-MAY-2002.
XX
PF      29-OCT-2001; 2001WO-GB004789.
XX
PR      27-OCT-2000; 2000GB-00026333.
PR      24-NOV-2000; 2000GB-00028727.
PR      07-MAR-2001; 2001GB-00005640.
XX
PA      (CHIR-) CHIRON SPA.
PA      (GENO-) INST GENOMIC RES.
XX
XX      Telford J, Masignani V, Margarit Y Rosi, Grandi G, Fraser C;
PI      Tettelin H;
XX
XX      WPI; 2002-352536/38.
DR      N-PSDB; ABN68156.
XX
XX      New Streptococcus protein for the treatment or prevention of infection or
PT      disease caused by Streptococcus bacteria, such as meningitis, and for
PT      detecting a compound that binds to the protein.
XX
XX      Claim 1; Page 3580; 4525pp; English.
PS
XX
XX      The invention relates to a protein (ABP25413-ABP30895) from group B
CC      streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC      (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC      the specification. The proteins have antibacterial and antiinflammatory
CC      activity. (i), nucleic acids encoding (i), ABN6044-ABN71526 and
CC      antibodies that bind (i) are used in the manufacture of medicaments for
CC      the treatment or prevention of infection or disease caused by
CC      Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC      Nucleic acids encoding (i) are used to detect Streptococcus in a
```

```
CC      biological sample. (I) is used to determine whether a compound binds to
CC      (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC      used as a vaccine or diagnostic composition. The disease caused by
CC      Streptococcus that is prevented or treated may be meningitis. Nucleic
CC      acid encoding (I) may be used to recombinantly produce (I) and may be
CC      used in gene therapy. Antibodies to (I) are used for affinity
CC      chromatography, immunoassays, and distinguishing/identifying
CC      Streptococcus proteins
XX
SQ      Sequence 376 AA;
Query Match      12.5%; Score 136; DB 5; Length 376;
Best Local Similarity 23.9%; Pred. No. 2.1e-05;
Matches 47; Conservative 36; Mismatches 90; Indels 24; Gaps 6;

QY      6 DLNIDMTVMPLQKSHLRDLKELLSYMPHEICFSDFIGEEGFLVDRP-----DN 57
DB      153 NISDILYALPGQTMDDVRNVAKALSINPHLSLSYLILEHHTVFMNKRGRKHLHLPTE 212

QY      58 SIDSEKLMWFCALCELSNGVINYETNFALKGHESRHNKLNWELKPHLGLGLYAVSLFLC 117
DB      213 DLEAEMFEYIIE-MERNGFEHYEISNFTKPGFESRHNLMYWDNVEYGVGAGASGYL-- 269

QY      118 NDKNNVRALIRKTGSF-----VKANNHLVTPELLEDEFFVYHFIQGLGTIQGVSLRA 171
DB      270 ----DGIR--YRNRGPIQHYLKGVSNGNARLSEVLKSNEMEEELFLGLRKKEGVSIK 323

QY      172 LRLRFEYN-EKQFFQFI 187
DB      324 FEQKFGTSFEKRYGQIV 340

RESULT 15
AEA49154
ID      AEA49154 standard; protein; 381 AA.
XX
AC      AEA49154;
XX
XX
DT      25-AUG-2005 (first entry)
XX
DE      L. rhamnosus polypeptide #18.
XX
XX      Gene fusion; wound healing; immunostimulant; antilipemic; vulnerary.
XX
OS      Lactobacillus rhamnosus.
XX
XX      WO2005056801-A1.
XX
XX      23-JUN-2005.
XX
XX      15-DEC-2003; 2003WO-NZ000278.
XX
XX      15-DEC-2003; 2003WO-NZ000278.
XX      (GENE-) GENESIS RES & DEV CORP LTD.
XX      (FONT-) FONTERA COOP GROUP LTD.
XX
XX      Glenn M, Havukkala I, Lubbers MW, Dekker J;
PI      WPI; 2005-435589/44.
DR      N-PSDB; AEA49072.
XX
XX      Isolated Lactobacillus polynucleotide, useful for modulating activity of
PT      polypeptide in a cell, and for production of more effective probiotic
PT      bacteria.
XX
XX      Claim 18; SEQ ID NO 98; 323pp; English.
PS
XX
XX      The invention relates to an isolated Lactobacillus rhamnosus
CC      polynucleotide and the polypeptide it encodes. The invention also relates
CC      to a genetic construct comprising the polynucleotide, a transgenic host
CC      cell comprising the genetic construct, a transgenic organism comprising
CC      the host cell or its progeny and a method of modulating the activity of a
```


This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:28:54 ; Search time 7.73918 Seconds
(without alignments)
2536.214 Million cell updates/sec

Title: US-10-688-058-10
Perfect score: 1090
Sequence: 1 RKPPFDLNDMTVMPLQKK.....QFINYCSTLSKKFVDDNIM 204
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1090	100.0	377	2 G70181	oxygen-independent
2	205	18.8	374	2 G97057	coproporphyrinogen
3	175.5	16.1	396	2 AD2425	coproporphyrinogen
4	151.5	13.9	374	2 F89939	hypothetical prote
5	148.5	13.6	370	2 H70423	oxygen-independent
6	141.5	13.0	385	2 G83817	coproporphyrinogen
7	138.5	12.7	374	2 D72288	hypothetical prote
8	137.5	12.6	379	2 C86767	oxygen-independent
9	136	12.5	393	2 D71354	probable oxygen-in
10	134.5	12.3	378	2 B71564	probable copropor
11	134	12.3	376	2 A98030	coproporphyrinogen
12	134	12.3	376	2 B95164	hypothetical prote
13	130.5	12.0	399	2 C71728	probable oxygen-in
14	126.5	11.6	412	1 S75358	coproporphyrinogen
15	124.5	11.4	366	1 B69640	coproporphyrinogen
16	124	11.4	391	2 G82320	probable oxygen-in
17	122.5	11.2	400	2 C97727	hypothetical prote
18	117	10.7	345	2 E71843	oxygen-independent
19	116	10.6	376	2 AD0116	probable oxygen-in
20	114.5	10.5	385	2 AD1259	coproporphyrinogen
21	114.5	10.5	385	2 AH1621	coproporphyrinogen
22	113	10.4	352	1 B64673	coproporphyrinogen
23	107	9.8	383	2 B82673	conserved hypothet
24	106	9.7	399	2 AE2616	hypothetical prote
25	106	9.7	402	2 E97398	probable oxygen-in
26	105.5	9.7	388	2 B75557	probable copropor
27	105	9.6	384	2 A83597	hypothetical prote
28	104.5	9.6	376	2 B84994	coproporphyrinogen
29	102.5	9.4	373	2 B86538	coproporphyrinogen

coproporphyrinogen
probable oxidoredu
probable oxygen-in
hypothetical prote
probable oxidase I
probable oxidase I
probable oxidase Y
hypothetical prote
coproporphyrinogen
coproporphyrinogen
coproporphyrinogen
flagellar hook-len
coproporphyrinogen
cyanodine receptor
DNA primase (EC 2.
hypothetical prote
multiple banded an
glycosyltransferas
probable membrane

ALIGNMENTS

RESULT 1

G70181

oxygen-independent coproporphyrinogen III oxidase homolog - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C;Accession: G70181

R;Praser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White,
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997

A;Authors: Smith, H.O.; Venter, J.C.

A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A;Reference number: A70100; MUID:98065943; PMID:9403685

A;Accession: G70181

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-377 <KLE>

A;Cross-references: UNIPROT:O51600; UNIPARC:UPI0000057576; GB:AE001167; GB:AE000783; NID

A;Experimental source: strain B31

Query Match 100.0%; Score 1090; DB 2; Length 377;
Best Local Similarity 100.0%; Pred. No. 6.1e-88;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	RKPPFDLNDMTVMPLQKKSHLRDLKELLSYMPHEICFSDPICBEEGVLRDFDONSID	60
DB	141	RKPPFDLNDMTVMPLQKKSHLRDLKELLSYMPHEICFSDPICBEEGVLRDFDONSID	200
QY	61	SEKLWFCALCELESNGYINYEITNFALKGESHNNKLNWELKPHGLGLYAVSLPCNDK	120
DB	201	SEKLWFCALCELESNGYINYEITNFALKGESHNNKLNWELKPHGLGLYAVSLPCNDK	260
QY	121	NNNVRALIRKTSFVANNHLVTFELLEDFEPVYHFIOGLGTIOGVSRLALRLFEYNE	180
DB	261	NNNVRALIRKTSFVANNHLVTFELLEDFEPVYHFIOGLGTIOGVSRLALRLFEYNE	320
QY	181	KQFFQFINYCSTLSKKFVDDNIM	204
DB	321	KQFFQFINYCSTLSKKFVDDNIM	344

RESULT 2

G97057

coproporphyrinogen III oxidase [imported] - Clostridium acetobutylicum

C;Species: Clostridium acetobutylicum

C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004

C;Accession: G97057

R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A;Reference number: A96900; MUID:21359325; PMID:21359325

A;Accession: G97057
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-374 <KUR>
A;Cross-references: UNIPROT:Q97J5; UNIPARC:UPI000000CA157; GB:AE001437; PIDN:AAK79250.1;
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC1279
C;Superfamily: oxygen-independent coproporphyrinogen oxidase

Query Match 18.8%; Score 205; DB 2; Length 374;
Best Local Similarity 32.4%; Pred. No. 2e-10;
Matches 62; Conservative 32; Mismatches 77; Indels 22; Gaps 8;

Qy 6 DLNIDMTVNNPQKKSHLKRDLKELLSYMPHEHCFSDFICEE-----EGFVLRDFDN 57
Db 151 NINVDLMFGIPDQDTLEDPKESLEFITKLKPEHISYSLIVEEGTPYFKMNEGKJK-LPN 209
Qy 58 SIDSEKLFWCALECSNGYINVEITNFALKGHESRHNKLNWELKPHLGLGLYAVSLIFC 117
Db 210 EDERDMYSFARTFLEKGNQVEISNFAVKDECRNLIYWELDNVYIGCGSAHS--YF 267
Qy 118 ND---KN--NNVALIRK--TGSFVKANNHLVTFELLEDFEFFVYHFIOGLGTIQGVSLRA 171
Db 268 NGVRYRINNKKYIEQISKGNSVVERNRL--LKEDMEEFNF---LGLRKTGRGVSIEE 322
Qy 172 LRLRFEYNEKQFP 184
Db 323 FKLKFNKDIOEVY 335

RESULT 3
AD2425
coproporphyrinogen III oxidase [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AD2425
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AD2425
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-396 <KUR>
A;Cross-references: UNIPROT:Q8YMH7; UNIPARC:UPI000000CECED; GB:BA000019; PIDN:BAB76655.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: ali4956
C;Superfamily: oxygen-independent coproporphyrinogen oxidase

Query Match 16.1%; Score 175.5; DB 2; Length 396;
Best Local Similarity 26.3%; Pred. No. 8.4e-08;
Matches 52; Conservative 38; Mismatches 87; Indels 21; Gaps 6;

Qy 6 DLNIDMTVNNPQKKSHLKRDLKELLSYMPHEHCFSDFICEE---GFVLRDFDNDISE 62
Db 167 EFSIDLSGLPHQSLDQWSDLTAVNIAPTHISIDLTTPGTAFGRYVKPGDNPPLPTD 226
Qy 63 ----KLWFCALCECSNGYINVEITNFALKGHESRHNKLNWELKPHLGLGLYAVSLIFCN 118
Db 227 ETTVKMYQMOKILITGGDYEHYEISNVAKPQCHRNRYWENPYPGFGNGAASYV--- 283

Qy 119 DKNNNVALIRKTSFVK-----ANNHLVTFELLEDFEFFVYHFIOGLGTIQGVSLRAL 172
Db 284 EGKFTRP--RKTKEYYQWQELIANHGVIDWEITPRADVLTLETMLGLRLADGVSLAAL 341

Qy 173 LRLRFEYNEKQFPFINYC 190
Db 342 TEEF---GKEKIQELHQC 356

RESULT 4
F89939
hypothetical protein hemN [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: F89939
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguci
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaico, C.; Sekimizu, K.; I
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayaashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: F89939
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-374 <KUR>
A;Cross-references: UNIPROT:Q99TR6; UNIPARC:UPI000000D767E; GB:BA000018; PID:gl3701381; P:
A;Experimental source: strain N315
C;Genetics:
A;Gene: hemN
C;Superfamily: oxygen-independent coproporphyrinogen oxidase

Query Match 13.9%; Score 151.5; DB 2; Length 374;
Best Local Similarity 26.3%; Pred. No. 1e-05;
Matches 56; Conservative 44; Mismatches 84; Indels 29; Gaps 9;

Qy 7 LNIDMTVNNPQKKSHLKRDLKELLSYMPHEHCFSDFICE-----EEGFVLRDFD 56
Db 150 ISLDLMYHLPKQTIEDFEQSLDLDMDIQHISYGLILEPKTQFYNNYRKGLLKLUPND 209
Qy 57 NSIDSEKLFWCALECSNGYINVEITNFALKGHESRHNKLNWELKPHLGLGLYA----- 111
Db 210 LGADMVQL---LMSKIEQSPFFHOYEISNFDLGDHSEHNKYVWFNEYIYFGAGASYVD 266
Qy 112 -VSLLEFCNKNRVALIRKTSFVKANNHLVTFELLEDFEFFVYHFIOGLGTIQGVSLR 170
Db 267 GVRTYTNIPNVHYIKAINKESKAILVSNKPSLT-ERMEE-EMFL-----GLRLNEGVSS 319

Qy 171 ALRLREYN-EKOPFOFINVCSTLSKKFVFDN 202
Db 320 RFKKFQDSIESVFGQTIN--NLKEKELIVEKN 350

RESULT 5
H70423
oxygen-independent coproporphyrinogen III oxidase - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: H70423
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: H70423
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-370 <AQF>
A;Cross-references: UNIPROT:O67418; UNIPARC:UPI0000056603; GB:AE000739; NID:G2983813; PI
A;Experimental source: strain VF5
C;Genetics:
A;Gene: hemF
C;Superfamily: oxygen-independent coproporphyrinogen oxidase

Query Match 13.6%; Score 148.5; DB 2; Length 370;
Best Local Similarity 26.6%; Pred. No. 1.8e-05;
Matches 51; Conservative 39; Mismatches 83; Indels 19; Gaps 6;

Qy 6 DLNIDMTVNNPQKKSHLKRDLKELLSYMPHEHCFSDFICEE-----GFVLRDFDND 60
Db 152 NVNVDLTWGWPTQSLKDLVELEHLKFPVKHVSAYLLTLYEDTPMGLAFKRGSEKPLEE 211

they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A>Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770; PMID:9665876
A:Accession: D71354
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-393 <COL>
A:Cross-references: UNIPROT:O83216; UNIPARC:UPI00000D3217; GB:AE001202; GB:AE000520; NID
A:Experimental source: strain Nichols
C:Genetics:
A:Gene: TP0186
C:Superfamily: oxygen-independent coproporphyrinogen oxidase

Query Match 12.5%; Score 136; DB 2; Length 393;
Best Local Similarity 25.6%; Pred. No. 0.00024;
Matches 50; Conservative 26; Mismatches 73; Indels 46; Gaps 8;

Qy 7 LNIDMTVMPLQKKSHLKRDLKELLSYMPHICFSDPIC-----EEEGFVLRFDFNSI 59
Db 168 ISADLIAGLRGQTARMVREDIDELLSFGLRHVSLYG-LCVPHPTTQERIA----- 218

Qy 60 DSEKLFCALECLSENGYINYEITNFA-LKGHESRHNKLNWELKPHGLGLGYAVSLFCN 118
Db 219 ---ALWAHGSAYLVRAGFNRYELSNFARTAADESAAHNRAYRWAPHAGVGPAGVGTFRVN 275

Qy 119 DKNNN-----VRALIRKTGSFVKANNHL-----VTLELLEDFEFVYH-----FIQGL 161
Db 276 LSLSKEGAWAIRSVRK-----HLGQYLAECVRENVEYHEFLTEHMCVQALLMGL 326

Qy 162 GTIQGVSLRALRLRF 176
Db 327 RLEQGLDVVTFARP 341

RESULT 10
B71564
Probable coproporphyrinogen III oxidase - Chlamydia trachomatis (serotype D, strain UW3/
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C:Accession: B71564
R:Stephens, R.S.; Kalman, S.; Lammell, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998
A>Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trac
A:Reference number: A71570; MUID:99000809; PMID:9784136
A:Accession: B71564
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-378 <ARN>
A:Cross-references: UNIPROT:O84055; UNIPARC:UPI00000D32E8; GB:AE001279; GB:AE001273; NID
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: hemN_1
C:Superfamily: oxygen-independent coproporphyrinogen oxidase

Query Match 12.3%; Score 134.5; DB 2; Length 378;
Best Local Similarity 27.9%; Pred. No. 0.00031;
Matches 43; Conservative 26; Mismatches 70; Indels 15; Gaps 4;

Qy 6 DLNIDMTVMPLQKKSHLKRDLKELLSYMPHICFSDPICEEGFVL-----RDFDNDSDS 61
Db 155 NVSADLIYGLTQTSDFIVDLHQALSLPIQHISLYNTIDPHTSFYKHKRILPSIADD 214

Qy 62 EKLMFCALECLSENGYINYEITNFA-LKGHESRHNKLNWELKPHGLGLGYAVSLFCN 118
Db 215 DSLAEMALABELLENQGFTRFELASYAKNOAASKHNTYYTAKPFLGLGVSASQYL--- 271

Qy 119 DKNNNVA--LIRKTGSFVKANNHLVTLELLEDL 150
Db 272 ---HGIRSKNISRISHYLRARAHQHLPTLESMEEL 302

RESULT 11

A98030
coproporphyrinogen oxidase (EC 1.3.3.3) [imported] - Streptococcus pneumoniae (strain R6;
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: A98030
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A>Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: A98030
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-376 <KUR>
A:Cross-references: UNIPROT:Q97Q22; UNIPARC:UPI0000051830; GB:AE007317;
C:Genetics:
A:Gene: hemN
C:Superfamily: oxygen-independent coproporphyrinogen oxidase
C:Keywords: oxidoreductase

Query Match 12.3%; Score 134; DB 2; Length 376;
Best Local Similarity 26.4%; Pred. No. 0.00034;
Matches 51; Conservative 28; Mismatches 90; Indels 24; Gaps 6;

Qy 5 FD-LNIDMTVMPLQKKSHLKRDLKELLSYMPHICFSDPICEEGFVLRFD----- 55
Db 151 FDNISIDLIYALPGQTMQEVKNVAKAIGLDIPHMSLYSLILENHTVPMNRRGKPLP 210

Qy 56 DNSIDSEKLFCALECLSENGYINYEITNFA-LKGHESRHNKLNWELKPHGLGLGYAVSL 115
Db 211 KEELEAEMFEYIIAE-LERAGFEHYELSNFKGFSRHNLMYWDNAEYYIGAGASGYV 269

Qy 116 FCNDKNNVR-----ALIRKTGSFVKANNHLVTLELLEDFEFVYHFIQGLGTIOGVSLRA 171
Db 270 -----NGVRYKNRGPTRHYLSAVEGNACITDHLKSQKEBEMFLGLRKKSGVSM-- 321

Qy 172 LRLRFEYNEKQFF 184
Db 322 --ARFEKFGQSF 332

RESULT 12
B95164
hypothetical protein SPI409 [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C:Accession: B95164
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A>Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: B95164
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-376 <KUR>
A:Cross-references: UNIPROT:Q97Q22; UNIPARC:UPI0000051830; GB:AE005672;
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SPI409
C:Superfamily: oxygen-independent coproporphyrinogen oxidase

Query Match 12.3%; Score 134; DB 2; Length 376;
Best Local Similarity 26.4%; Pred. No. 0.00034;
Matches 51; Conservative 28; Mismatches 90; Indels 24; Gaps 6;

Qy 5 FD-LNIDMTVMPLQKKSHLKRDLKELLSYMPHICFSDPICEEGFVLRFD----- 55
Db 151 FDNISIDLIYALPGQTMQEVKNVAKAIGLDIPHMSLYSLILENHTVPMNRRGKPLP 210

Search completed: January 24, 2006, 19:59:03
Job time : 8.73918 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:27:53 ; Search time 43.8352 Seconds
(without alignments)
3283.383 Million cell updates/sec

Title: US-10-688-058-10
Perfect score: 1090
Sequence: 1 RKPPFDLNDVTNMPLOKK.....QFINYCSTLSKKVFDDNIM 204

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1090	100.0	377	051600 BORBU	051600 borrelia bu
2	934	85.7	377	Q66014 BORGA	Q66014 borrelia ga
3	210	19.3	365	Q8R695 FUSNN	Q8R695 fusobacteri
4	205.5	18.9	375	Q8RB71 THETN	Q8RB71 thermoanaer
5	205	18.8	374	Q97JJ5 CLOAB	Q97JJ5 clostridium
6	192.5	17.7	382	Q992Q7 CLOTE	Q992Q7 clostridium
7	190.5	17.5	165	Q7P440 FUSNV	Q7P440 fusobacteri
8	189	17.3	377	Q8X188 CLOPE	Q8X188 clostridium
9	175.5	16.1	396	Q8YMH7 ANASP	Q8YMH7 anabaena sp
10	175	16.1	408	Q7NC94 GLOVI	Q7NC94 gloeobacter
11	172.5	15.8	391	Q73MP2 TREDE	Q73MP2 treponema d
12	163.5	15.0	375	Q8KAY8 CHLZE	Q8KAY8 chlorobium
13	161.5	14.8	374	Q6GG87 STAAR	Q6GG87 staphylococ
14	157.5	14.4	374	Y498_BUCBP	Q89A47 buchnera ap
15	154	14.1	378	Q5LZC1 STRT1	Q5LZC1 streptococc
16	151.5	13.9	374	Q6G8V4 STAA5	Q6G8V4 staphylococ
17	151.5	13.9	374	Q7A0S0 STAAW	Q7A0S0 staphylococ
18	151.5	13.9	374	Q7ASC2 STAAW	Q7ASC2 staphylococ
19	151.5	13.9	374	Q99TR6 STAAW	Q99TR6 staphylococ
20	151.5	13.9	385	Q3HFH7 STAAC	Q3HFH7 staphylococ
21	151	13.9	378	Q5M3X3 STRT2	Q5M3X3 streptococc
22	150	13.8	374	Q9RPI0 9FIRM	Q9RPI0 desulfitoba
23	150	13.8	375	Q5X2T4 LEGPA	Q5X2T4 legionella
24	150	13.8	375	Q5ZT11 LEGPH	Q5ZT11 legionella
25	150	13.8	376	Q7MMW0 PORGI	Q7MMW0 porphyromon
26	149.5	13.7	371	Q7VIJ0 HELHP	Q7VIJ0 helicobacte
27	148.5	13.6	370	Q67418 AQUAE	Q67418 aquifex aeo
28	146	13.4	375	Q5WU92 LEGPL	Q5WU92 legionella
29	145.5	13.3	374	Q4FF16_9THEM	Q4FF16 thermotoga
30	145	13.3	408	Q7UTS9_PROMP	Q7UTS9 prochloroco
31	142.5	13.1	351	Q8EVQ8_MYCPE	Q8EVQ8 mycoplasma

32	142.5	13.1	383	2	05QY62 IDILO	05QY62 idiomarina
33	141.5	13.0	376	2	Q99ZW6_STRP8	Q99ZW6 streptococc
34	141.5	13.0	376	2	Q8PI77_STRP8	Q8PI77 streptococc
35	141.5	13.0	385	2	Q9KD75_BACHD	Q9KD75 bacillus ha
36	141.5	13.0	398	2	Q878L1_STRP3	Q878L1 streptococc
37	141.5	13.0	408	2	Q8K7R5_STRP3	Q8K7R5 streptococc
38	141.5	13.0	408	2	Q5XCD8_STRP6	Q5XCD8 streptococc
39	141	12.9	349	2	Q4HKR7_CAMLA	Q4HKR7 campylobact
40	140	12.8	378	2	Q81LR9_BACAN	Q81LR9 bacillus an
41	140	12.8	379	2	Q6HT74_BACAN	Q6HT74 bacillus an
42	140	12.8	379	2	Q634M4_BACCC	Q634M4 bacillus ce
43	139.5	12.8	383	2	Q93R30_TETHA	Q93R30 tetragenoco
44	139	12.8	379	2	Q5KWZ4_GEOKA	Q5KWZ4 geobacillus
45	138.5	12.7	374	2	Q9X0P9_THEMEA	Q9X0P9 thermotoga

ALIGNMENTS

RESULT 1

ID	051600 BORBU	PRELIMINARY;	PRT;	377 AA.
AC	051600;			
DT	01-JUN-1998	(TrEMBLrel. 06, Created)		
DT	01-JUN-1998	(TrEMBLrel. 06, Last sequence update)		
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)		
DE	Oxygen-independent coproporphyrinogen III oxidase, putative.			
GN	OrderedLocusNames=BB0656;			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OC	Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;			
OC	Borrelia burgdorferi group.			
OX	NCBI_TaxID=139;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=ATCC 35210 / B31;			
EX	MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;			
RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,			
RA	Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,			
RA	Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,			
RA	Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,			
RA	Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,			
RA	Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A.,			
RA	Artiach P., Bowman C., Gariand S.A., Fujii C., Cotton M.D., Horst K.,			
RT	"Genomic sequence of a Lyme disease spirochaete, Borrelia			
RT	burgdorferi."			
RL	Nature 390:580-586(1997).			
DR	EMBL; AE001167; AAC67009.1; -; Genomic_DNA.			
DR	PIR; G70181; G70181.			
DR	TIGR; BB0656; -			
DR	GO; GO:0003824; F:catalytic activity; IEA.			
DR	GO; GO:0005506; F:iron ion binding; IEA.			
DR	InterPro; IPR006638; ELP3/MiaB/NifB.			
DR	InterPro; IPR007197; Radical_SAM.			
DR	Pfam; PF04055; Radical_SAM; I.			
DR	SMART; SM00729; ELP3; I.			
KW	Complete proteome.			
SQ	SEQUENCE 377 AA; 44756 MW; B07859CFC7793561 CRC64;			

Query Match 100.0%; Score 1090; DB 2; Length 377;
Best Local Similarity 100.0%; Pred. No. 3.9e-87;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	RKPPFDLNDVTNMPLOKKSHLKRDLKELLSYMPHEICFSDFICEEGFVLRDFNSID	60
Db	141	RKPPFDLNDVTNMPLOKKSHLKRDLKELLSYMPHEICFSDFICEEGFVLRDFNSID	200
Qy	61	SEKLWFCALCELSNGYINYEITNFALKGHESRHNKLNWELKPHLGLGLVAVSLFCNDK	120
Db	201	SEKLWFCALCELSNGYINYEITNFALKGHESRHNKLNWELKPHLGLGLVAVSLFCNDK	260
Qy	121	NNNVRALIRKTSFVKANNHLVTFELLEDLEFFVYHFIOGLGTIOGVSILRALRPFYNE	180


```
Db 261 NNNVRALIRKTSFVKANNHLVTFELLEDLEFFVYHFHFIQGLGTIGQVSLRALRLRFYNE 320
Qy 181 KQPFQFNYCSTLSKKFVFDNIM 204
Db 321 KQPFQFNYCSTLSKKFVFDNIM 344

RESULT 2
Q660L4_BORGA
ID Q660L4_BORGA PRELIMINARY; PRT; 377 AA.
AC Q660L4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE HemN-related protein.
GN OrderedLocusNames=BG0679;
OS Borrelia garinii.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=29519;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBI;
RA Gloeckner G., Lehmann R., Romualdi A., Pradella S.,
RA Schulte-Spechtel U., Wilske B., Suehnel J., Platzer M.;
RT "Comparative analysis of the Borrelia garinii genome.";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000013; AAU07507.1; -; Genomic_DNA.
DR InterPro; IPR007197; Radical SAM.
DR Pfam; PF04055; Radical SAM; 1.
DR SMART; SM00729; Elp3; 1.
DR Complete proteome.
KW SEQUENCE 377 AA; 44968 MW; C8192B7F0CD24051 CRC64;

Query Match 85.7%; Score 934; DB 2; Length 377;
Best Local Similarity 84.8%; Pred. No. 1.8e-73;
Matches 173; Conservative 14; Mismatches 17; Indels 0; Gaps 0;

Qy 1 RKFPFDLNDMTVMNPLOKSHLRDLKELLSYMPHEHICFSDFTCEEGFVLRDFNSID 60
Db 141 RKFSFDLNDMTINIPQKSDLRDLKELLSYMPHEHICFSDFTCEEGFVLRDFNSID 200

Qy 61 SEKLWFCALCELSNGYINYEITPALKGHESRNKLNWELKPHGLGLYAVSLFCNDK 120
Db 201 SEKLWFALEYLSNGYINYEITPALKGHESRNKLNWELKPHGLGLYAVSLFCNDK 260

Qy 121 NNNVRALIRKTSFVKANNHLVTFELLEDLEFFVYHFHFIQGLGTIGQVSLRALRLRFYNE 180
Db 261 NNNVRALIRKDSGFVKASNLAKFELLEDLEFFYHFHFIQGLGTIGQVNLRLRFYND 320

Qy 181 KQPFQFNYCSTLSKKFVFDNIM 204
Db 321 KQPFHFINCYNLKSKFVFDNIM 344

RESULT 3
Q8R695_FUSNN
ID Q8R695_FUSNN PRELIMINARY; PRT; 365 AA.
AC Q8R695;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Oxygen-independent coproporphyrinogen III oxidase (EC 1.-.-.-).
GN OrderedLocusNames=FN0560;
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
```

```
RX DOI=10.1128/JB.184.7.2005-2018.2002;
RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fonstein M., Kyrpides N.C., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586.";
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL; AE009951; AAL94756.1; -; Genomic_DNA.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR GO; GO:0004109; F: coproporphyrinogen oxidase activity; IEA.
DR GO; GO:0005506; F: iron ion binding; IEA.
DR GO; GO:0006779; P: porphyrin biosynthesis; IEA.
DR InterPro; IPR006638; Elp3/MiAB/NifB.
DR InterPro; IPR010723; HemN C.
DR InterPro; IPR004559; HemN rel.
DR InterPro; IPR007197; Radical SAM.
DR Pfam; PF06969; HemN C; 1.
DR Pfam; PF04055; Radical SAM; 1.
DR SMART; SM00729; Elp3; 1.
DR TIGRFAMs; TIGR00539; HemN_rel; 1.
DR Complete proteome; Oxidoreductase.
KW SEQUENCE 365 AA; 43029 MW; 41EEF343A04309D6 CRC64;

Query Match 19.3%; Score 210; DB 2; Length 365;
Best Local Similarity 32.0%; Pred. No. 5.1e-10;
Matches 56; Conservative 32; Mismatches 61; Indels 26; Gaps 5;

Qy 1 RKFPFD-LNIDMTVMNPLOKSHLRDLKELLSYMPHEHICFSDFTCEEGFVLRDF--- 55
Db 147 RECQFDNISLDIMFSLPYQTLMLQNDLEKLVSLNPNHISYLSIWEEGTKFFRDLKSGK 206

Qy 56 -----DNSIDSEKLWFCALCELSNGYINYEITPALKGHESRNKLNWELKPHGLGLY 111
Db 207 LKETDNDLEA-SMYEYIIIEFLKSDYIHEISNFKSKDFESRNSTYWNKKYLGVLGA 265

Qy 112 VSLI-----FCNDKNNVRALIRK---TGSFVKANNHLVTFELLEDL 150
Db 266 AGYLNVRVYKNFNLKDYNNLDRNLIPDEKEILTEEDIEQRYLVGFRLLNKI 320

RESULT 4
Q8R71_THETN
ID Q8R71_THETN PRELIMINARY; PRT; 375 AA.
AC Q8R71;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Coproporphyrinogen III oxidase and related FeS oxidoreductases.
GN Name=HemN; OrderedLocusNames=TT0952;
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336; DOI=10.1101/gr.219302;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of the T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
DR EMBL; AE013060; AAM24208.1; -; Genomic_DNA.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0004109; F: coproporphyrinogen oxidase activity; IEA.
DR GO; GO:0005506; F: iron ion binding; IEA.
DR GO; GO:0006287; F: magnesium ion binding; IEA.
DR GO; GO:0016491; F: oxidoreductase activity; IEA.
DR GO; GO:0006779; P: porphyrin biosynthesis; IEA.
DR InterPro; IPR002453; Beta_tubulin.
```


[illegible]

```
Qy 51 VLRFDFNSIDSEKLMFCALECLSGYINVEITNPFALKGHESRHNKLWELKPHLGLGLY 110
Db 212 L--NLPNEDIERMYRVAIFLENGHGHQHEISNFSKREHCINLIYWDLSEYIGGCLA 269
Qy 111 AVSLL----FCNDKNNVRALIRKTSFVKANNHL-VTFELLEDLEFFVYHFYHFIQGLGTIQ 165
Db 270 AHSPLKGYRYSNVHNIEDIYKLIENKNKIKINTYKNTKDTMEEFM-----MGLRKIK 323
Qy 166 GVSRLRALRPFYNEKQFFQFYNVCSLTKKFFV 198
Db 324 GINTEEFYKGFPHKNIEVY-----GDIIRKYYI 350

RESULT 7
Q7P440_FUSNV
ID Q7P440_FUSNV PRELIMINARY; PRT; 165 AA.
AC Q7P440;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Coproporphyrinogen oxidase, anaerobic (EC 1.-).
GN Name=FNV0357;
OS Fusobacterium nucleatum subsp. vincentii ATCC 49256.
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=209882;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 49256;
RA Karpatrial V., Ivanova N., Anderson I., Reznik G., Bhattacharyya A.,
RA Gardner W L., Mikhailova N., Larsen N., D'Souza M., Walunas T.,
RA Haselkorn R., Overbeek R., Kyrpides N.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAF01000140; EAA23389.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0004109; F:coproporphyrinogen oxidase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006779; P:porphyrin biosynthesis; IEA.
DR InterPro; IPR010723; HemN_C.
DR Pfam; PF06969; HemN_C; 1.
KW Oxidoreductase.
SQ SEQUENCE 165 AA; 19650 MW; C4532CBBC40DAE01 CRC64;

Query Match 17.5%; Score 190.5; DB 2; Length 165;
Best Local Similarity 31.1%; Pred. No. 1.1e-08;
Matches 50; Conservative 30; Mismatches 56; Indels 25; Gaps 4;

Qy 14 NMPIQKSHLRDLKELLSYMPHEHICFSDPICEBEGFVLDRF-----DNSIDSEKLM 65
Db 3 SLPNQTWAMLQNDLEKLVNLPNPHISYLSIWEETKFFDLKSGKLKETDNDLEA-SMY 61
Qy 66 FCALECLSGYINVEITNPFALKGHESRHNKLWELKPHLGLGLYAVSLL----- 115
Db 62 EYIIEFLSKDYTHYEISNFSKKGESSHNSIYENKNYLVGSLSAAGYLDNVRYKNFFN 121
Qy 116 ---FCNDKNNVRALIRK---TGSFVKANNHLVTFELLEDL 150
Db 122 LKDYNNDKKNLIPIDEXEILTQEDIEQRYLVGFRLNKI 162

RESULT 8
Q8XIS8_CLOPE
ID Q8XIS8_CLOPE PRELIMINARY; PRT; 377 AA.
AC Q8XIS8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Probable coproporphyrinogen III oxidase.
GN Name=hemN; OrderedLocusNames=CPE2036;
```

```
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=13 / Type A;
RC MEDLINE=21664373; PubMed=11792842; DOI=10.1073/pnas.022493799;
RA Shimizu T., Ohtani K., Hirakawa H., Ohehima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RA "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001 (2002).
DR EMBL; BA00016; BAB81742.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004109; F:coproporphyrinogen oxidase activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0006779; P:porphyrin biosynthesis; IEA.
DR InterPro; IPR006638; Elp3/MiaB/NifB.
DR InterPro; IPR010723; HemN_C.
DR InterPro; IPR004559; HemN_rel.
DR InterPro; IPR007197; Radical_SAM.
DR Pfam; PF06969; HemN_C; 1.
DR Pfam; PF04055; Radical_SAM; 1.
DR SMART; SM00729; Elp3; 1.
DR TIGRFAMs; TIGR00539; hemN_rel; 1.
KW Complete proteome.
SQ SEQUENCE 377 AA; 43713 MW; 0BCEFAD409B34F78 CRC64;

Query Match 17.3%; Score 189; DB 2; Length 377;
Best Local Similarity 26.5%; Pred. No. 3.7e-08;
Matches 49; Conservative 37; Mismatches 83; Indels 16; Gaps 4;

Qy 6 DLNIDMTVMNPLQKSHLRDLKELLSYMPHEHICFSDPICEBEGFVLDRDN-----S 58
Db 154 NINVDLMFGLPNQRLNWEETLREIISLEPAHISAYSLSLIEEGTAFYNLYENDKLPTE 213
Qy 59 IDSEKLMFCALECLSGYINVEITNPFALKGHESRHNKLWELKPHLGLGLYAVSLLFCN 118
Db 214 EERKMVHLAKKILEENGFNQYEISNYSYKGEKGRHNLAYNMNDNWIWGVSAAS--YIN 271
Qy 119 DKN-----NNVRALIRKTSFVKANNHLVTFELLEDLEFFVYHFYHFIQGLSRLRL 174
Db 272 GKRIKNISSVEKYINSINEKREAVEEIIINNSKNDMEEFM---MGLRKINGIDENEFKN 328
Qy 175 RFEYN 179
Db 329 RFSMN 333

RESULT 9
Q8YMH7_ANASP
ID Q8YMH7_ANASP PRELIMINARY; PRT; 396 AA.
AC Q8YMH7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Coproporphyrinogen III oxidase.
GN OrderedLocusNames=all4956;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120.";
```

DR Pfam; PF04055; Radical_SAM; 1.
DR TIGRFAMs; TIGR00539; hemN_rel; 1.
DR Complete proteome.
SQ SEQUENCE 408 AA; 45119 MW; 936C6CD8C5992C39 CRC64;

Query Match 16.1%; Score 175; DB 2; Length 408;
Best Local Similarity 27.9%; Pred. No. 6.8e-07;
Matches 53; Conservative 27; Mismatches 76; Indels 34; Gaps 5;

Qy 6 DLNIDMTVMNPLQCKSHLKRDLKELLSYMPHEITCFSDFICEREGFVLRF-----DNS 58
Db 187 NFNLDLIFGLPHQSLADWEFSEVLRLFTHLSDLYDLILEETPFGRFLGGQAPLPGD 246
Qy 59 IDEKULWFCALCELENGYINYEITPALKGHESRNKLNWELKPHLGLGLYAVSLFLCN 118
Db 247 DDTVMYLRACERLQCGYRQYEIANFARPGFCRHRVYVWDRNPYYGIG-----296
Qy 119 DKNNVRLIRKTSFVKANNHLVT--FELLEDEF-----FVHFIOGLGTIQ 166
Db 297 ---NGATGYVR--GARVERPRLLDHYFAMVEGGEFFAGAPVMPEELADTLILGLRLEG 351
Qy 167 VLSRALRLRF 176
Db 352 VEVNALVQRF 361

RESULT 11
Q73MP2_TREDE
ID Q73MP2_TREDE PRELIMINARY; PRT; 391 AA.
AC Q73MP2;
DT 05-JUL-2004 (TEMBLrel. 27, Created)
DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE Oxygen-independent coporphyrinogen III oxidase, putative.
GN OrderedLocusNames=TDE1466;
OS Treponema denticola.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=158;
RN [1]
NN NUCLEOTIDE SEQUENCE.
RC STRATN=ATCC 35405 / DSM 14222;
RK PubMed=15064399; DOI=10.1073/pnas.0307639101;
RA Seshadri R., Myers G.S.A., Tettelin H., Eisen J.A., Heidelberg J.F.,
RA Dodson R.J., Davidson T.M., DeBoy R.T., Fouts D.E., Haft D.H.,
RA Selengut J., Ren Q., Brinkac L.M., Madupu R., Kolonay J.F.,
RA Durkin S.A., Daugherty S.C., Shetty J., Shvartsbeyn A.,
RA Gebrgeorgis E., Geer K., Tsegaye G., Malek J.A., Ayodeji B.,
RA Shatsman S., McLeod M.P., Smajs D., Howell J.K., Pal S., Amin A.,
RA Vashisth P., Neilell T.Z., Xiang Q., Sodergren E., Baca E.,
RA Weinstein G.M., Norris S.J., Fraser C.M., Paulsen I.T.;
RT "Comparison of the genome of the oral pathogen Treponema denticola
RT with other spirochete genomes";
RL Proc. Natl. Acad. Sci. U.S.A. 101:5646-5651(2004).
DR EMBL; AE017251; AAS11983.1; -; Genomic_DNA.
DR TIGR; TDE1466; -
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0004109; F:coporphyrinogen oxidase activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006779; P:porphyrin biosynthesis; IEA.
DR InterPro; IPR006638; Elp3/Miab/NifB.
DR InterPro; IPR010723; HemN C.
DR InterPro; IPR004559; HemN_rel.
DR InterPro; IPR007197; Radical_SAM.
DR Pfam; PF06969; HemN C; 1.
DR Pfam; PF04055; Radical_SAM; 1.
DR SMART; SM00729; Elp3; 1.
DR TIGRFAMs; TIGR00539; hemN_rel; 1.
DR Complete proteome.
SQ SEQUENCE 391 AA; 45034 MW; 8D121969939FF1CB CRC64;

Query Match 15.8%; Score 172.5; DB 2; Length 391;

Best Local Similarity 26.9%; Pred. No. 1.1e-06;
Matches 50; Conservative 39; Mismatches 72; Indels 25; Gaps 7;

Qy 7 LNTDMVTNMPLOKSHLKRDLKELLSYMPHE-----ICFSDPFCIEERGFLVRDPDINSIDS 61
Db 162 LSCDLIAGLNKQSLGILKDDIQTULHPKPEHFSYALCSNTKUPPEK-----DDEI-- 212
Qy 62 EKLWFCALCELESNGYINYEITNFALKG-HESRHNKLNWELKPHLGLGLYAVSLFCNDK 120
Db 213 AELWSYCKDILEKSNYKVEFSYKLNYSIHNEKYWRLEDYIGVGPAFGSIFPDRK 272
Qy 121 ----NNVRAL-----IKRTGSF--VKANNHLVTFLELLEFFVYHFIQGLGTIQGVSLR 170
Db 273 VLHANAHANALRFAIKNEIKMTLKNRDEVYEVETIDEKEFIEBAFMMGLRLTEGIDRS 332
Qy 171 ALRLRF 176
Db 333 FFKSRF 338

RESULT 12
Q8KAY8 CHLTE PRELIMINARY; PRT; 375 AA.
AC Q8KAY8;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Oxygen-independent coproporphyrinogen III oxidase, putative (EC 1.-.-
DE -).
GN OrderedLocusNames=CT2010;
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobaculum.
OX NCBI_TaxID=1097;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TLS / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=12093901; DOI=10.1073/pnas.132181499;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA Dodson R.J., Deboy R.T., Gwinn M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.F., Yang F.,
RA Holt I.E., Unayam L.A., Mason T.M., Brenner M., Shea T.P.,
RA Parksey D.S., Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B.,
RA Radune D., Vamathavan J.J., Khouri H.M., White O., Gruber T.M.,
RA Ketchum K.A., Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
RT photosynthetic, anaerobic, green-sulfur bacterium.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
DR EMBL; AE006470; AAM73228.1; -; Genomic_DNA.
DR TIGR; CT2010; -;
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004109; F:coproporphyrinogen oxidase activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0006779; P:porphyrin biosynthesis; IEA.
DR InterPro; IPR006638; ELP3/MiaB/NifB.
DR InterPro; IPR010723; HemN_C.
DR InterPro; IPR004559; HemN_rel.
DR Pfam; PF04055; Radical_SAM.
DR SMART; SM00729; Elp3; I.
DR TIGRFAMs; TIGR00539; hemN_rel; 1.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 375 AA; 41628 MW; 1772A85DCF885C20 CRC64;

Query Match 15.0%; Score 163.5; DB 2; Length 375;
Best Local Similarity 32.6%; Pred. No. 6.3e-06;
Matches 45; Conservative 24; Mismatches 58; Indels 11; Gaps 3;

Qy 2 KPFPDLNIDMTVMNPLQKSHLKRDLKELLSYMPHEICFSDPFCIEERGFLV-----LRD 54
Db 148 RFP-SVSDILMCGAEGEMLAEWEGDLRAALDLCPQHISYIMLTLEKTRLWRVVRKGLRD 206

Qy 55 FDNISDEKLFWFCALCELESNGYINYEITNFALKGHESRHNKLNWELKPHLGLGLYAVSL 114
Db 207 LPCEEAQAAMRYMAAELLGNAGYGHYEISNFALEGHHSRYNLASWMEPYLGFQPAHSF 266
Qy 115 LFCND---KNNNVRALIR 129
Db 267 LVDGDAETRFNSVSSLTR 284

RESULT 13
Q6GB7 STAAR PRELIMINARY; PRT; 374 AA.
AC Q6GB7;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Putative oxygen-independent coproporphyrinogen III oxidase.
GN OrderedLocusNames=SAR1660;
OS Staphylococcus aureus (strain MRSA252).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282458;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC PubMed=15213324; DOI=10.1073/pnas.04025211101;
RA Holden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Fellwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitsch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL; BX571856; CAG40655.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0004109; F:coproporphyrinogen oxidase activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006779; P:porphyrin biosynthesis; IEA.
DR InterPro; IPR006638; ELP3/MiaB/NifB.
DR InterPro; IPR010723; HemN_C.
DR InterPro; IPR004559; HemN_rel.
DR Pfam; PF06969; HemN_C; 1.
DR SMART; SM00729; Elp3; I.
DR TIGRFAMs; TIGR00539; hemN_rel; 1.
KW Complete proteome.
SQ SEQUENCE 374 AA; 42894 MW; 2C43A6C957BA81DA CRC64;

Query Match 14.8%; Score 161.5; DB 2; Length 374;
Best Local Similarity 27.2%; Pred. No. 9.4e-06;
Matches 58; Conservative 43; Mismatches 83; Indels 29; Gaps 9;

Qy 7 LNIDMTVMNPLQKSHLKRDLKELLSYMPHEICFSDPFCIE-----BEGFV-LRDFD 56
Db 150 ISLDLMYHLPKQTTEDFEQSLDLALDMDIQHISYGLILEPKTQFYNNYRKGLKLPNED 209
Qy 57 NSIDSEKLFWFCALCELESNGYINYEITNFALKGHESRHNKLNWELKPHLGLGLYA----- 111
Db 210 LGADMYQL---LMSKIEQSPFHQYEISNFALDGHESEHNKYTFNFEYFGAGASGYVD 266
Qy 112 -VSLFLCNDNNVRALIRKTSFGVKANNHLVTFLELLEFFVYHFIQGLGTIQGVSLR 170
Db 267 GVRTNINPNVHYKAINKESKAILVSNKPSLT-ERNEE-EMFL-----GLRLNEGYSR 319
Qy 171 ALRLRFPYN-EKQFPQFIFYNCSTLSKCFVFDN 202

Db 320 RFKKKFDQSIESTVFGQTIN--NLKEKKLIVEKN 350

RESULT 14

```
Y498 BUCBP
ID Y498 BUCBP STANDARD; PRT; 374 AA.
AC Q89A47;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical protein bhp498.
GN OrderedLocusNames=bhp498;
OS Buchnera aphidicola (subsp. Baizongia pistaciae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=135842;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=22436901; PubMed=12522265; DOI=10.1073/pnas.0235981100;
RA van Ham R.C.H.J., Kamerbeek J., Palacios C., Raueell C., Abascal F.,
RA Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,
RA Tamames J., Viguera E., Latorre A., Valencia A., Moran F., Moya A.;
RT "Reductive genome evolution in Buchnera aphidicola.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).
CC -I- SIMILARITY: Belongs to the anaerobic coproporphyrinogen III
CC oxidase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AB014017; AAC27203.1; -; Genomic_DNA.
CC InterPro; IPR010723; HemN_C.
CC InterPro; IPR004559; HemN_rel.
CC InterPro; IPR007197; Radical_SAM.
CC Pfam; PF06969; HemN_C; 1.
CC Pfam; PF04055; Radical_SAM; 1.
CC TIGRFAMs; TIGR00539; hemN_rel; 1.
CC Complete proteome; Hypothetical protein; Oxidoreductase;
CC Porphyrin biosynthesis.
CC SEQUENCE 374 AA; 44752 MW; 9D93C81AD2565053 CRC64;
Query Match 14.4%; Score 157.5; DB 1; Length 374;
Best Local Similarity 22.5%; Pred. No. 2.1e-05;
Matches 45; Conservative 35; Mismatches 73; Indels 47; Gaps 6;
Qy 6 DLNIDMTVNNPLQKSHLRDLKELLSYMPHEHICFSDPFICEEGFVLRFDFNSIDSEKJW 65
Db 156 NINLDLYGLPKQSLQEQALLDLKTAISLKPNIHSWCFEYIEKNNNYKNLSKSCNLIW 215
Qy 66 FCALE---CLSENGYINVEITNFKLGHSRHNKLNWELKPHLGILGYAVSLFPCNDKN 122
Db 216 KIFLQGEKLLKSKYKYSYSYKNTYQCLNLYNWKFGDYLIG-----CNAHGK 267
Qy 123 NVRALIRKTSFVKA-----NNHLVTFELLEDLEFFVYHFIOGLGTIQG 166
Db 268 -----ITQNGKIITKKNKLNKFPWNGKYTYKNHISKNL-SLEFF----- 309
Qy 167 VSLRALRLRFEYNEKQFFQF 186
Db 310 --NNRLRLNTPIYRKDFKY 327
```

RESULT 15

```
QSLZC1 STRT1 PRELIMINARY; PRT; 378 AA.
ID QSLZC1 STRT1 PRELIMINARY; PRT; 378 AA.
AC QSLZC1;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Coproporphyrinogen III oxidase.
```

```
GN Names=hemN; OrderedLocusNames=stri1236;
OS Streptococcus thermophilus (strain CNRZ 1066).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=299768;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinkis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulakauskas S., Lapidus A., Goltzman E., Mazur M., Pusch G.D.,
RA Fonstein M., Overbeek R., Kyprides N., Purnelle B., Prozzi D.,
RA Ngui K., Masuy D., Hancy F., Burtan S., Boutry M., Delcour J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
RT bacterium Streptococcus thermophilus.";
RL Nat. Biotechnol. 22:1554-1558(2004).
DR EMBL; CP000024; AAV62781.1; -; Genomic_DNA.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0004109; F:coproporphyrinogen oxidase activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006779; P:porphyrin biosynthesis; IEA.
DR InterPro; IPR006638; ELP3/MiaB/NifB.
DR InterPro; IPR010723; HemN_C.
DR InterPro; IPR004559; HemN_rel.
DR InterPro; IPR007197; Radical_SAM.
DR Pfam; PF06969; HemN_C; 1.
DR Pfam; PF04055; Radical_SAM; 1.
DR SMART; SM00729; ELP3; 1.
DR TIGRFAMs; TIGR00539; hemN_rel; 1.
KW Complete proteome.
SQ SEQUENCE 378 AA; 43487 MW; 450DD78F16B0D966 CRC64;
Query Match 14.1%; Score 154; DB 2; Length 378;
Best Local Similarity 26.9%; Pred. No. 4.3e-05;
Matches 59; Conservative 33; Mismatches 83; Indels 42; Gaps 10;
Qy 6 DLNIDMTVNNPLQKSHLRDLKELLSYMPHEHICFSDPFICEEGFVLRFDF-----DN 57
Db 153 NISIDLIYALPGQTMQVKNVRKALELDIPHLISLISLLEHHTVFNKMRGKLNLPTE 212
Qy 58 SIDSEKLNWFCALCLESNGYINVEITNFKLGHSRHNKLNWELKPHLGILGYAVSLFPC 117
Db 213 DLEAEMFDYIIQF--LEKNGFHEYIEISNFTKPGMESRHNLMYNNDEYIGVGAGSGV-- 269
Qy 118 NDKNNNVRLIRKTG---SFVKA-----NNHLVTFELLEDLEFFVYHFIOGLGTI 164
Db 270 ----NGVR--YNRGPIQHYLKAIADGCHARLNEEHHTKVEEMEE-EFFL-----GLRKK 317
Qy 165 QGVSLRALRLRFEYNEKQFFQFINYCSTLSKQFVFD 200
Db 318 SGYSIK----RFB--EKFGLSFSDTYGDIVKQLQED 347
Search completed: January 24, 2006, 19:56:46
Job time : 45.8352 secs
```

This Page Blank (uspio)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	145	13.3	155	2	US-09-830-230A-522	Sequence 522, App
2	145	13.3	175	2	US-09-830-230A-521	Sequence 521, App
3	134	12.3	376	2	US-09-583-110-3497	Sequence 3497, Ap
4	134	12.3	409	2	US-09-107-433-4312	Sequence 4312, Ap
5	112	10.3	391	2	US-09-134-001C-3690	Sequence 3690, Ap
6	111	10.2	393	2	US-09-107-532A-6801	Sequence 6801, Ap
7	105.5	9.7	401	2	US-09-502-540-14907	Sequence 14907, A
8	105	9.6	424	2	US-09-252-991A-20695	Sequence 20695, A
9	104.5	9.6	397	2	US-09-134-000C-5885	Sequence 5885, Ap
10	102.5	9.4	394	2	US-09-198-452A-396	Sequence 396, App
11	102.5	9.4	402	2	US-09-438-185A-382	Sequence 382, App
12	98	9.0	387	2	US-09-489-039A-11948	Sequence 11948, A
13	96.5	8.9	382	2	US-09-543-681A-5208	Sequence 5208, Ap
14	93	8.5	5037	2	US-09-424-783-4	Sequence 4, Appli
15	88.5	8.1	287	2	US-09-830-230A-122	Sequence 122, App
16	88.5	8.1	309	2	US-09-830-230A-121	Sequence 121, App
17	88.5	8.1	387	2	US-09-248-796A-22214	Sequence 22214, A
18	88	8.1	5032	2	US-09-538-092-979	Sequence 979, App
19	86	7.9	389	2	US-09-328-352-8088	Sequence 8088, Ap
20	85	7.8	460	2	US-09-198-452A-959	Sequence 959, App
21	85	7.8	460	2	US-09-438-185A-891	Sequence 891, App
22	82	7.5	694	2	US-09-538-092-1351	Sequence 1351, Ap
23	82	7.5	698	2	US-09-949-016-10215	Sequence 10215, A
24	81.5	7.5	1255	2	US-09-248-796A-14158	Sequence 14158, A
25	81	7.4	254	2	US-09-134-000C-6525	Sequence 6525, Ap
26	81	7.4	268	2	US-09-543-681A-7944	Sequence 7944, Ap
27	81	7.4	899	2	US-09-248-796A-15467	Sequence 15467, A

```
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 521
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-830-230A-521

Query Match      13.3%; Score 145; DB 2; Length 175;
Best Local Similarity 90.0%; Pred. No. 2.3e-08;
Matches 27; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RKFPFDLNDMTVNMPLQKSHLRDLKEL 30
Db 141 RKFPFDLNDMTVNMPLQKSHLRDLQRI 170

RESULT 3
US-09-583-110-3497
; Sequence 3497, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/593,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3497
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3497

Query Match      12.3%; Score 134; DB 2; Length 376;
Best Local Similarity 26.4%; Pred. No. 1.2e-06;
Matches 51; Conservative 28; Mismatches 90; Indels 24; Gaps 6;

Qy 5 FD-LNIDMTVNMPLQKSHLRDLKELSYNPEHICFSDFICEBEGFVLDRF----- 55
Db 151 FDNISIDLIALPGQTMQVKNVAKAIGLDIPHMSLSYLLENHTVPMNMRGKPLP 210

Qy 56 DNSIDSEKLFWCALECLSNYINYEITNPAKKGHESRHNKLWELKPHLGLGLYAVSLL 115
Db 211 KEELEAEFMFYITAE--LERAGFEHYEISNFKPGFESRHNLMYWDNAEYYIGAGAGYV 269

Qy 116 FCNDKNNVNR---ALIRKTSFVKANNHLVTFLELLEDFVYHFIOGLGTIOQVSLRA 171
Db 270 -----NGVRYKHGPIRHYLSAVEEGNACITEDHLSQKEQMEEMFLGLRKKSGVSM-- 321

Qy 172 LRLRFEYNEKQFF 184
Db 322 --ARFEKFGQSF 332

RESULT 4
US-09-107-433-4312
; Sequence 4312, Application US/09107433
; Patent No. 6800744
);
```

```
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Walham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4312:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 409 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...409
; SEQUENCE DESCRIPTION: SEQ ID NO: 4312:
US-09-107-433-4312

Query Match      12.3%; Score 134; DB 2; Length 409;
Best Local Similarity 26.4%; Pred. No. 1.4e-06;
Matches 51; Conservative 28; Mismatches 90; Indels 24; Gaps 6;

Qy 5 FD-LNIDMTVNMPLQKSHLRDLKELSYNPEHICFSDFICEBEGFVLDRF----- 55
Db 184 FDNISIDLIALPGQTMQVKNVAKAIGLDIPHMSLSYLLENHTVPMNMRGKPLP 243

Qy 56 DNSIDSEKLFWCALECLSNYINYEITNPAKKGHESRHNKLWELKPHLGLGLYAVSLL 115
Db 244 KEELEAEFMFYITAE--LERAGFEHYEISNFKPGFESRHNLMYWDNAEYYIGAGAGYV 302

Qy 116 FCNDKNNVNR---ALIRKTSFVKANNHLVTFLELLEDFVYHFIOGLGTIOQVSLRA 171
Db 303 -----NGVRYKHGPIRHYLSAVEEGNACITEDHLSQKEQMEEMFLGLRKKSGVSM-- 354

Qy 172 LRLRFEYNEKQFF 184
Db 355 --ARFEKFGQSF 365

RESULT 5
US-09-134-001C-3690
; Sequence 3690, Application US/09134001C
; Patent No. 6380370
);
```



```

/ TELEPHONE: (781)893-5007
/ TELEFAX: (781)893-8277
/ INFORMATION FOR SEQ ID NO: 6801:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 393 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: YES
/ ORIGINAL SOURCE:
/ ORGANISM: Enterococcus faecium
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (B) LOCATION 1...393
/ SEQUENCE DESCRIPTION: SEQ ID NO: 6801:
US-09-107-532A-6801

Query Match 10.2%; Score 111; DB 2; Length 393;
Best Local Similarity 20.7%; Pred. No. 0.00054;
Matches 42; Conservative 37; Mismatches 88; Indels 36; Gaps 5;

QY 6 DLNIDMTVMPLQKKSHLKRDLKELLSYMPHEHICFSDFTCEESGFVLR-----DFDNS 58
DB 168 NVSIDLIYALPGGTLEGYLDLTKALALDLPVLSYLSILENKTMTMNVWVROGRLELPDQ 227
QY 59 IDSEKLFWCALECSNGYINVEITNPFALKGHESRNKLNWELKPHLGLGLYAVSLFCN 118
DB 228 EYETRMFEETIAMEKHGRHVEISNFGLSCHESQENLWYNNNDHYFGFGAGASYL--- 284
QY 119 DKNNVVALIRKTGSFVKANNHLVTPLELLEDEFFVYHFTQGLGTIOGVSLRALRL-RFE 177
DB 285 ---GDIR--YRNKGP-----IQHYLRPLRANERPVLBEKLRKN 319
QY 178 YNEKQFFQFYNVCSTLSKKFVED 200
DB 320 QIEEMFGLGRKKTGITKKHYD 342

RESULT 7
US-09-902-540-14907
/ Sequence 14907, Application US/09902540
/ Patent No. 6833447
/ GENERAL INFORMATION:
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Wiegand, Roger C.
/ TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
/ FILE REFERENCE: 38-10(15849)B
/ CURRENT APPLICATION NUMBER: US/09/902,540
/ CURRENT FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: 60/217,883
/ PRIOR FILING DATE: 2000-07-10
/ NUMBER OF SEQ ID NOS: 16825
/ SEQ ID NO 14907
/ LENGTH: 401
/ TYPE: PRT
/ ORGANISM: Myxococcus xanthus
US-09-902-540-14907

Query Match 9.7%; Score 105.5; DB 2; Length 401;
Best Local Similarity 23.2%; Pred. No. 0.0023;
Matches 47; Conservative 28; Mismatches 83; Indels 45; Gaps 7;

QY 3 FPDFLNIDMTVMPLQKKSHLKRDLKELLSYMPHEHICFSDFTCEESGFVLR----- 54
DB 170 FFPV-VAMDFIYGVHGTVQVAEADARRAVALAPEHL--SYALTVERDVLAEETPLSKRL 226
QY 55 -----FDNSIDSEKLFWCALECSNGYINVEITNPFALKGHESRNKLNWELKPHL 105
DB 227 KRGELELPDDTVDVMDARV---VREYIGAAGLHRYEVSNHAREGFSRNHALYWTGGEYL 283
QY 106 GLGLYAVSLFCNDDKN--NVRALTRKTSFVKANNHLVTPLELLEDEFFVYHFTQGLGTIOGVSLRALRL-RFE 153

```

```
Db 284 ALGVGATGMLSPQPHRYVNLRS-----AERYIVEAEACAGLPCEGREGPLGPPELF 333
Qy 154 VYHFIOGLGTITQGVSLRALRLRF 176
Db 334 AERLANGRLVSGVDWEAVCERY 356

RESULT 8
US-09-252-991A-20695
; Sequence 20695, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20695
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20695

Query Match 9.6%; Score 105; DB 2; Length 424;
Best Local Similarity 26.0%; Pred. No. 0.0029;
Matches 32; Conservative 18; Mismatches 69; Indels 4; Gaps 2;

Qy 1 RKFPFD-LNIDMTVMNPLQKSHLKRDLKELLSYMPHEICFSDFICEEGFVLRDPNSI 59
Db 198 RQAGFDNPLDLMHGLPQSLDDALDLRQALAPATHLSVQLTMEPTVFWSQPPPLP 257
Qy 60 DSEKLN----FCALECLSENGYINYEITNPAKLGHSRKNLWELKPHLGLGLYAVSLLF 116
Db 258 EDDLWDIQEAGQALLDHEGYQVETSAYARDGLHARHNLNYSFGDFLGAGAHAKLS 317
Qy 117 CND 119
Db 318 APD 320

RESULT 9
US-09-134-000C-5885
; Sequence 5885, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5885
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5885

Query Match 9.6%; Score 104.5; DB 2; Length 397;
Best Local Similarity 26.0%; Pred. No. 0.003;
Matches 33; Conservative 14; Mismatches 53; Indels 27; Gaps 2;

Qy 6 DLNIDMTVMNPLQKSHLKRDLKELLSYMPHEICFSDFICE-----EE 48
```

```
Db 170 NVSIDLIYALPGQTLESFRDTLTRALALDLPFHYSLYSLSILENKTMTFMNWMVRQGRLOLPEE 229
Qy 49 GFVLRDFDINSIDSEKLFWFCALCELESNGYINYEITNPAKLGHSRKNLWELKPHLGLG 108
Db 230 EIEAQMFDETI-----EAMEKKGRHQYEVSNFALTGKESQHNLAYWNNDHYYGFG 279
Qy 109 LYAVSLL 115
Db 280 AGASGYL 286

RESULT 10
US-09-198-452A-396
; Sequence 396, Application US/09198452A
; Patent No. 6552994
; GENERAL INFORMATION:
; APPLICANT: Griffois, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 396
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-396

Query Match 9.4%; Score 102.5; DB 2; Length 394;
Best Local Similarity 21.9%; Pred. No. 0.005;
Matches 49; Conservative 40; Mismatches 96; Indels 39; Gaps 8;

Qy 6 DLNIDMTVMNPLQKSHLKRDLKELLSYMPHEICFSDP-----ICEE 47
Db 175 NLSIDLIYGLPTQSLSEIFLSDLHQALTPTTHISLYNLTTIDPHTSFYKRRKILVPTIAQE 234
Qy 48 EGFVLRDFDINSIDSEKLFWFCALCELESNGYINYEITNPAKLGHSRKNLWELKPHLGL 107
Db 235 E--ILAEM--SLIAENL-----LLSQGFQRYELASYAKPDYPAKHNLYYTDRPFJGL 283
Qy 108 GLYAVSLLFCNDKNN-----VRALIRKTGSPVKANNHLVTFELLEDLEFFVVFHTQ 160
Db 284 GVSASQYLH-GERSKNYSHISHYLRA-VRKNLPQETSEILPKKERIKEALALRLLEG 341
Qy 161 LGTIQGVSLRALRLRFEYNEKQFFQFINYCSTLSKK-FVFDNI 203
Db 342 ADLAEFPSTLISMLTQDVKLQNLPSVHGQCCLALNRQGRLPDHTI 385

RESULT 11
US-09-438-185A-382
; Sequence 382, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 382
```

```
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: Cpn0380
US-09-438-185A-382

Query Match      9.4%; Score 102.5; DB 2; Length 402;
Best Local Similarity 21.9%; Pred. No. 0.0052;
Matches 49; Conservative 40; Mismatches 96; Indels 39; Gaps 8;

Qy 6 DLNIDMTVMPLQKSHLKDRLKELLSYMPHEHICFSDP-----ICEE 47
Db 183 NLSIDLYGLPTQSLFSLDHOALTLPITHISLYNLTDPTHSFYKHKILVPTIAQE 242
Qy 48 EGFVLRDPDINSIDSEKLFWCALECLSGYINYEITNFALKGHESRHNKLNWELKPHLGL 107
Db 243 E--LLAEM--SLLAENL-----LLSQGFQRYELASYKPDYPAKNLYYWDTPFLGL 291
Qy 108 GLYAVSLFLCNDKNN-----VRALIRKTSFVKANNHLVTFELLEDEFFVYHFIOG 160
Db 292 GVSASQYLH--GERSKNYSHISHYLRA--VRKNLPQTQETSEILPKKERIKALALRLLEG 349
Qy 161 LGTIQGVSLRALRLRPEYNEKQFPQFNYCSTLSKK--FVFPDNI 203
Db 350 ADLAEFPSTLISMLTQDVKLQNLFSVHGQCCLALNRQGLRPHDTI 393

RESULT 12
US-09-489-039A-11948
; Sequence 11948, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11948
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11948

Query Match      9.0%; Score 98; DB 2; Length 387;
Best Local Similarity 21.3%; Pred. No. 0.016;
Matches 36; Conservative 30; Mismatches 71; Indels 32; Gaps 6;

Qy 8 NIDMTVMPLQKSHLKDRLKELLSYMPHEHICFSDFICEEGFVLRDPDINSI----- 59
Db 169 NLDLMHGLPDQSLEALDRLQALNPPHLSYQLTIE-----PNTLFGSRPPVL 219
Qy 60 -DSEKLFWC---ALECLSGYINYEITNFALKGHESRHNKLNWELKPHLGLGLYAVSLL 115
Db 220 PDDDALDIFEGQHLLSAAAGTQYETSAYAKPGFQCOHNNLYWRFGLYIGCGAHGKI 279
Qy 116 FCND-----KNNVRLIRKTSFVKANNHLVTFELLEDEFFVYHF 157
Db 280 TFPDGRILRTAKTRHPRGYME--GYLSEHQHVE--EADKPFEPFMNRF 324

RESULT 13
US-09-543-681A-5208
; Sequence 5208, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
```

```
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; ORGANISM: Proteus mirabilis
; FEATURE:
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5208
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5208

Query Match      8.9%; Score 96.5; DB 2; Length 382;
Best Local Similarity 25.0%; Pred. No. 0.023;
Matches 26; Conservative 17; Mismatches 58; Indels 3; Gaps 1;

Qy 8 NIDMTVMPLQKSHLKDRLKELLSYMPHEHICFSDFICEEGFVLRDPDINSIDSEKLFWC 67
Db 167 NLDLMHGLPDQSLSQALSLSLQQAIALSPPHLSYVQLTIEPTQFGRPPKLPDDMLNDI 226
Qy 68 ALE---CLESGYINYEITNFALKGHESRHNKLNWELKPHLGLG 108
Db 227 FSEGDKLLTAAGYQQYETSAYCKPGFQCOHNNLYWRFGLYIG 270

RESULT 14
US-09-424-783-4
; Sequence 4, Application US/09424783
; Patent No. 6780608
; GENERAL INFORMATION:
; APPLICANT: Hakamata, Yasuhiro
; APPLICANT: Nishimura, Seiichi
; APPLICANT: Barsoumian, Edward Leon
; TITLE OF INVENTION: Human Type 3 Ryanodine Receptor Protein
; TITLE OF INVENTION: and DNA Molecules Coding Therefor
; FILE REFERENCE: 9652.2000000
; CURRENT APPLICATION NUMBER: US/09/424,783
; CURRENT FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: PCT/EP98/02926
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: DE 197 22 317.6
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 5037
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-424-783-4

Query Match      8.5%; Score 93; DB 2; Length 5037;
Best Local Similarity 24.0%; Pred. No. 2.3;
Matches 49; Conservative 36; Mismatches 75; Indels 44; Gaps 11;

Qy 13 VNNPLOKSHLKDRLK--ELLSYMPHEHICFSDFICEEGFV--LRDPDINSIDSEKLF 66
Db 4045 VDMLESSNVEMILKPFDMFLKLDIVGSEAFQDYVTDPRGLISKDFQKAMDQKQFT 4104
Qy 67 -----CALECLSGYINYEITNFALKGHESRHNKLNWELKPHLGLGLYAVSLLFCND 119
Db 4105 GPRIQFLLSCSEADENEMINFE--EFANRFQ-----PARDIGF----- 4141
Qy 120 KNNVRLIRKTSFVKANNHLVTFELLEDEFFVYHFVHFIQGLGTIQ--GVSRLRALRLRFE 177
Db 4142 ---NVAVLLTNLSEHVDPDLRNF--LELAESILEYFRPYLGRLEIMGASRIERYFE 4196
Qy 178 YNEKQFFQF--INYCSTLSKKFVFD 200
Db 4197 ISETNRAQWEMPQVKESKRFIFD 4220

RESULT 15
US-09-830-230A-122
```

```
; Sequence 122, Application US/09830230A
; Patent No. 6902893
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481US
; CURRENT APPLICATION NUMBER: US/09/830,230A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: PCT/US98/12718
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/057,483
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 122
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-830-230A-122

Query Match      8.1%; Score 88.5; DB 2; Length 287;
Best Local Similarity 22.7%; Pred. No. 0.13;
Matches 45; Conservative 25; Mismatches 67; Indels 61; Gaps 7;

Qy      3  PFPDLNIDMTVMPLQKKSHLKRDLKELLSYMPHEHICFSDPICEEGFVLRFDFNSIDSE 62
Db      14  FLFKIN-----KSLGK---KELPDQNTHCUSPEYDNLAKILIWDFKQLRKE 59

Qy      63  KLPFCALECLSGNYINYEITNFALKGHESRHKNLWELKPHLGLGLYAVSLFCNDKN 122
Db      60  -----GFFTQIKN-----DSSQYINARKN 79

Qy      123  NVRALIRKTS---FVKANHILVTFELLEDL-EFFVYHFIQGLGTIQGVSLRALRLRFEY 178
Db      80  NISFSIKREGSKITFECFNNHLI---IIQDLFRETLNLEKITKEVETVSLRAKCLDYSI 136

Qy      179  NEKQFFQFINYCSTLSKK 196
Db      137  NYDKILSNINLNKRIKKE 154
```

Search completed: January 24, 2006, 20:02:41
Job time : 13.5762 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:57:15 ; Search time 35.8542 Seconds
(without alignments)
2377.329 Million cell updates/sec

Title: US-10-688-058-10
Perfect score: 1090
Sequence: 1 RKFPFDLNDMTVMPLQKK.....QFINYCSTLSKKFVDDNIM 204

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_5/ptodata/1/pubpaa/US07_PUBCOMB.pdp.*
 - 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pdp.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pdp.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pdp.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pdp.*
 - 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1090	100.0	204	5	US-10-688-058-10
2	1090	100.0	377	5	US-10-688-058-12
3	151.5	13.9	385	5	US-10-470-048B-106
4	145	13.3	155	5	US-10-994-726-522
5	145	13.3	175	5	US-10-994-726-521
6	135	12.4	381	5	US-10-650-274-98
7	134	12.3	376	5	US-10-472-928-2860
8	134	12.3	409	5	US-10-617-320-4312
9	117	10.7	345	4	US-10-335-977-6717
10	117	10.7	352	4	US-10-335-977-6718
11	117	10.7	364	4	US-10-335-977-6719
12	104.5	9.6	493	4	US-10-425-115-261185
13	102.5	9.4	394	4	US-10-289-762-396
14	97	8.9	548	4	US-10-724-972A-7488
15	93	8.5	350	6	US-11-097-143-13632
16	91.5	8.4	605	4	US-10-282-122A-54636
17	91	8.3	1103	4	US-10-282-122A-76866
18	89.5	8.2	4473	4	US-10-437-963-201113
19	89.5	8.2	4737	4	US-10-437-963-201116
20	89	8.2	665	4	US-10-437-963-172583
21	88.5	8.1	287	5	US-10-994-726-122
22	88.5	8.1	309	5	US-10-994-726-121
23	88.5	8.1	701	4	US-10-282-122A-47140
24	88	8.1	1032	4	US-10-607-631-4
25	88	8.1	5032	4	US-10-408-765A-26
26	88	8.1	5038	5	US-10-723-860-1310
27	88	8.1	5038	5	US-10-756-149-5039

28	88	8.1	5081	4	US-10-276-774-1850	Sequence 1850, Ap
29	88	8.1	5081	5	US-10-450-763-35328	Sequence 35328, A
30	87	8.0	605	4	US-10-238-075-691	Sequence 691, App
31	85.5	7.8	463	6	US-11-097-143-14868	Sequence 14868, A
32	85	7.8	460	4	US-10-289-762-959	Sequence 959, App
33	84.5	7.8	204	4	US-10-335-977-7414	Sequence 7414, Ap
34	84.5	7.8	742	4	US-10-437-963-109290	Sequence 109290,
35	82.5	7.6	330	3	US-09-738-626-6014	Sequence 6014, Ap
36	82.5	7.6	410	4	US-10-156-761-13098	Sequence 13098, A
37	82.5	7.6	580	6	US-11-097-143-13371	Sequence 13371, A
38	82.5	7.6	739	4	US-10-437-963-155552	Sequence 155552,
39	82	7.5	483	4	US-10-437-963-125107	Sequence 125107,
40	82	7.5	694	3	US-09-842-758-75	Sequence 75, Appl
41	82	7.5	694	3	US-09-855-828-14	Sequence 14, Appl
42	82	7.5	694	4	US-10-345-680-26	Sequence 26, Appl
43	82	7.5	694	4	US-10-174-333-75	Sequence 75, Appl
44	82	7.5	694	5	US-10-978-282-11	Sequence 11, Appl
45	82	7.5	1067	4	US-10-437-963-121534	Sequence 121534,

ALIGNMENTS

RESULT 1

US-10-688-058-10
; Sequence 10, Application US/10688058
; Publication No. US20050058661A1
; GENERAL INFORMATION:
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: HALE, KATHERINE S.
; APPLICANT: JOHNSTON, STEPHEN A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF THE GENUS
; TITLE OF INVENTION: BORRELIA
; FILE REFERENCE: MCRO:003US
; CURRENT APPLICATION NUMBER: US/10/688,058
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 10
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-688-058-10

Query Match	100.0%;	Score 1090;	DB 5;	Length 204;
Best Local Similarity	100.0%;	Pred. No. 9.4e-106;		
Matches 204;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	RKFPFDLNDMTVMPLQKKSHLKRDLKELLSYMPHICFSDPFCBEEGVFLRDFDMSID	60	
Db	1	RKFPFDLNDMTVMPLQKKSHLKRDLKELLSYMPHICFSDPFCBEEGVFLRDFDMSID	60	
Qy	61	SEKLWCALECLSSNGYINYEITNFALKGHESHKLNWELKPHLGLYAVSLFCNDK	120	
Db	61	SEKLWCALECLSSNGYINYEITNFALKGHESHKLNWELKPHLGLYAVSLFCNDK	120	
Qy	121	NNNVRALIRKTSFVKANNHLVTFELLEDEFFVYHFIOQLGTIQGVSLRALRLRFEYNE	180	
Db	121	NNNVRALIRKTSFVKANNHLVTFELLEDEFFVYHFIOQLGTIQGVSLRALRLRFEYNE	180	
Qy	181	KQFFQFINYCSTLSKKFVDDNIM	204	
Db	181	KQFFQFINYCSTLSKKFVDDNIM	204	

RESULT 2

US-10-688-058-12
; Sequence 12, Application US/10688058
; Publication No. US20050058661A1
; GENERAL INFORMATION:
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: HALE, KATHERINE S.

```

; APPLICANT: JOHNSTON, STEPHEN A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF THE GENUS
; TITLE OF INVENTION: BORRELIA
; FILE REFERENCE: MCR0:0030US
; CURRENT APPLICATION NUMBER: US/10/688,058
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-688-058-12

Query Match      100.0%; Score 1090; DB 5; Length 377;
Best Local Similarity 100.0%; Pred. No. 2.1e-105;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 RKPPFDLNDMTVNMPLQKSHLKDRLKELLSYMPHEICFSDFTCEBEGFVLRDPFNSID 60
Db      141 RKPPFDLNDMTVNMPLQKSHLKDRLKELLSYMPHEICFSDFTCEBEGFVLRDPFNSID 200

Qy      61 SEKLWFCALCELENGYINYEITNFALKGHESRHNKLNWELKPHLGLGLYAVSLFCNDK 120
Db      201 SEKLWFCALCELENGYINYEITNFALKGHESRHNKLNWELKPHLGLGLYAVSLFCNDK 260

Qy      121 NNNVALIRKTGTFVKANNHLVTPLELLEDFVYHFHFIQGLGTIGQVSLRALRFRFEYNE 180
Db      261 NNNVALIRKTGTFVKANNHLVTPLELLEDFVYHFHFIQGLGTIGQVSLRALRFRFEYNE 320

Qy      181 KQFFQFNYCSTLKKFVDDNIM 204
Db      321 KQFFQFNYCSTLKKFVDDNIM 344

RESULT 3
US-10-470-048B-106
; Sequence 106, Application US/10470048B
; Publication No. US20050037444A1
; GENERAL INFORMATION:
; APPLICANT: MEINKE ET AL.
; TITLE OF INVENTION: A METHOD FOR IDENTIFICATION, ISOLATION AND PRODUCTION OF
; TITLE OF INVENTION: ANTIGENS TO A SPECIFIC PATHOGEN
; FILE REFERENCE: SONN:0350US
; CURRENT APPLICATION NUMBER: US/10/470,048B
; CURRENT FILING DATE: 2003-07-25
; NUMBER OF SEQ ID NOS: 603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 106
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-470-048B-106

Query Match      13.9%; Score 151.5; DB 5; Length 385;
Best Local Similarity 26.3%; Pred. No. 5.2e-07;
Matches 56; Conservative 44; Mismatches 84; Indels 29; Gaps 9;

Qy      7 LNIDMTVNMPLQKSHLKDRLKELLSYMPHEICFSDFTCEBEGFVLRDPFNSID 56
Db      161 ISLDLMLYHLPKQTTFDQSLDLDLMDIQHISYGLILEPTQFYNNYRKGLLKLPNED 220

Qy      57 NSIDSEKLWFCALCELENGYINYEITNFALKGHESRHNKLNWELKPHLGLGLYAVSLFCNDK 111
Db      221 LGADMYQL---LMSKIEQSPHQVEISNFDLGHESRHNKLNWELKPHLGLGLYAVSLFCNDK 277

Qy      112 -VSLLFPCNDKNNVALIRKTGTFVKANNHLVTPLELLEDFVYHFHFIQGLGTIGQVSLR 170
Db      278 GVRVYTNINPVRHYTKAINKESKAILVSNKPSLT-ERMEE-EMFL-----GLRLNEGVSSS 330

Qy      171 ALRLRFEYN-EKQFFQFNYCSTLKKFVDDNIM 202

```

```

Db      331 RFKKKFDQSIESVFGQTIN--NLKKEKELIVEKN 361

RESULT 4
US-10-994-726-522
; Sequence 522, Application US/10994726
; Publication No. US20050147999A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481D1
; CURRENT APPLICATION NUMBER: US/10/994,726
; CURRENT FILING DATE: 2004-11-23
; PRIOR APPLICATION NUMBER: 09/830,230
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: PCT/US98/12718
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/057,483
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 522
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-994-726-522

Query Match      13.3%; Score 145; DB 5; Length 155;
Best Local Similarity 90.0%; Pred. No. 7.5e-07;
Matches 27; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 RKPPFDLNDMTVNMPLQKSHLKDRLKEL 30
Db      121 RKPPFDLNDMTVNMPLQKSHLKDRLQRI 150

RESULT 5
US-10-994-726-521
; Sequence 521, Application US/10994726
; Publication No. US20050147999A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481D1
; CURRENT APPLICATION NUMBER: US/10/994,726
; CURRENT FILING DATE: 2004-11-23
; PRIOR APPLICATION NUMBER: 09/830,230
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: PCT/US98/12718
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/057,483
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 521
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-994-726-521

Query Match      13.3%; Score 145; DB 5; Length 175;

```

Best Local Similarity 90.0%; Pred. No. 8.8e-07;
Matches 27; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RKFPFDLNDMTVMNPLQKKSHLKRDLKEL 30
Db 141 RKFPFDLNDMTVMNPLQKKSHLKRDLQRI 170

RESULT 6
US-10-650-274-98
; Sequence 98, Application US/10650274
; Publication No. US20050202437A1
; GENERAL INFORMATION:
; APPLICANT: GLENN, MATTHEW
; APPLICANT: HAVUKKALA, ILKKA J
; APPLICANT: LUBBERS, MARK WILLIAM
; APPLICANT: DEKKER, JAMES
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES,
; TITLE OF INVENTION: MATERIALS INCORPORATING THEM, AND METHODS FOR USING
; TITLE OF INVENTION: THEM.
; FILE REFERENCE: 11000.1073
; CURRENT APPLICATION NUMBER: US/10/650.274
; CURRENT FILING DATE: 2003-08-28
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: FASTSEQ FOR WINDOWS VERSION 4.0
; SEQ ID NO 98
; LENGTH: 381
; TYPE: PRT
; ORGANISM: LACTOBACILLUS RHAMNOSUS
US-10-650-274-98

Query Match 12.4%; Score 135; DB 5; Length 381;
Best Local Similarity 26.2%; Pred. No. 2.8e-05;
Matches 53; Conservative 33; Mismatches 76; Indels 40; Gaps 8;

Qy 1 RKFPFDLNDMTVMNPLQKKSHLKRDLKELSYMPHEICFSDPFCIEEGFVLDRDINSI 59
Db 147 RKVGFNLSIDLFLPDPQSRDDFLNSLQKALADLPHYSTYSILERKTI----FYNLM 202

Qy 60 DSEKL-----WFCALCELESNGYINYEITNFALKGHESRHNKLNWELKPHLGLG 108
Db 203 RQRKLRLPTQDVEADMYQDAIDLMEAGHGHQYELSNFAKTYQCRHNLLYWQNDKYFGG 262

Qy 109 LYAVSLFLCNDKNNVRALIRKTSF-----VKANNHLVTFLELLEDFPVVYHFIQGLG 162
Db 263 AGAFGYL-GRDZYNYGPIKQYLAPLHADHPLVLAHLVFPVSEQIEB-EMEL-----GLR 315

Qy 163 TIQGVSLRALRPEYNEKQFF 184
Db 316 TMAGV-----NEDRFY 326

RESULT 7
US-10-472-928-2860
; Sequence 2860, Application US/10472928
; Publication No. US20050020813A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926W0
; CURRENT APPLICATION NUMBER: US/10/472.928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 2860
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: oxygen-independent coproporphyrinogen III oxidase, putative

; OTHER INFORMATION: Cellular location: cytoplasm
; OTHER INFORMATION: Similar to strain R6 sequence 15903309 (0.E+01)
US-10-472-928-2860

Query Match 12.3%; Score 134; DB 5; Length 376;
Best Local Similarity 26.4%; Pred. No. 3.4e-05;
Matches 51; Conservative 28; Mismatches 90; Indels 24; Gaps 6;

Qy 5 FD-LNIDMTVMNPLQKKSHLKRDLKELLSYMPHEICFSDPFCIEEGFVLDRDF-----55
Db 151 FDNISIDLIALPQQTMEQVKENVAKAIGLIDIPHMSLYSLILENHTVFMNMRMRGKLP 210

Qy 56 DNSIDSEKLFWCALECLNGYINYEITNFALKGHESRHNKLNWELKPHLGLGLYAVSL 115
Db 211 KEELEAEFFYIIAE-LERAGFHEIYELSNFKPFSRHNMTWDNAEYIGIGAGASYV 269

Qy 116 FCNDKNNVR-----ALIRKTSFVKANNHLVTFLELLEDFVYHFIQGLGTQGVSLRA 171
Db 270 -----NGVYKNGHPIRHYLSAVESGNACITEDHLSQKEQMEEMFLGLRKKSQVSM--321

Qy 172 LRLRPEYNEKQFF 184
Db 322 --ARFEKFGQSF 332

RESULT 8
US-10-617-320-4312
; Sequence 4312, Application US/10617320
; Publication No. US20050136404A1
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS

NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/617,320
FILING DATE: 10-Jul-2003
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
APPLICATION NUMBER: 60/085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 4312:
SEQUENCE CHARACTERISTICS:
LENGTH: 409 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae

US-10-335-977-6718

Query Match 10.7%; Score 117; DB 4; Length 352;
Best Local Similarity 23.3%; Pred. No. 0.0019;
Matches 45; Conservative 29; Mismatches 51; Indels 68; Gaps 9;
Qy 6 DLNIDMTVNPLOKKSHLKRDLEKLLSYMP-EHICFSDFCIEEGFVLR-----DFD 56
Db 157 NISIDLIYNTPLDNETSLKEELK-LAKELPTNHL SAYALSIKNTNLEKNAKPSVD 215
Qy 57 NSIDSEKLMFCALECLSGYINYEITNFALKGHESRHNKLNWELKPHLGLG----- 108
Db 216 NVV-----REVLEGFQKQYEVSNYA-RNYQVKHNLAYWGAKDYLCGAGAVCVA 265
Qy 109 ---LYAVSL-----FCNDK-----NNVRLALIR 129
Db 266 NERFYAKKLIENYIKDPLKRQVETLNKQDKLEKFLGLRCELGVLSLLDENKVKFLIE 325
Qy 130 KTGSFVKANNHLV 142
Db 326 ENKAFIK-NNRLI 337

RESULT 11

US-10-335-977-6719
; Sequence 6719, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 6719:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8) LOCATION 1...364
; SEQUENCE DESCRIPTION: SEQ ID NO: 6719:

US-10-335-977-6719

Query Match

Best Local Similarity 10.7%; Score 117; DB 4; Length 364;
Matches 45; Conservative 29; Mismatches 51; Indels 68; Gaps 9;
Qy 6 DLNIDMTVNPLOKKSHLKRDLEKLLSYMP-EHICFSDFCIEEGFVLR-----DFD 56
Db 169 NISIDLIYNTPLDNETSLKEELK-LAKELPTNHL SAYALSIKNTNLEKNAKPSVD 227
Qy 57 NSIDSEKLMFCALECLSGYINYEITNFALKGHESRHNKLNWELKPHLGLG----- 108
Db 228 NVV-----REVLEGFQKQYEVSNYA-RNYQVKHNLAYWGAKDYLCGAGAVCVA 277
Qy 109 ---LYAVSL-----FCNDK-----NNVRLALIR 129
Db 278 NERFYAKKLIENYIKDPLKRQVETLNKQDKLEKFLGLRCELGVLSLLDENKVKFLIE 337
Qy 130 KTGSFVKANNHLV 142
Db 338 ENKAFIK-NNRLI 349

RESULT 12

US-10-425-115-261185
; Sequence 261185, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 261185
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_169815C.1.pep
US-10-425-115-261185

Query Match 9.6%; Score 104.5; DB 4; Length 493;

Best Local Similarity 24.8%; Pred. No. 0.061;
Matches 31; Conservative 25; Mismatches 56; Indels 13; Gaps 3;
Qy 8 NIDMTVNPLOKKSHLKRDLEKLLSYMP-EHICFSDFCIEE-----EGFVLRDFDSDSE 62
Db 240 SMDLISSLPNQTEEMWEEELRCTVDARPTHSVVDLQIEQGTGFGQWYTPGVFPLPSDTE 299
Qy 63 KLWF--CALECLSGYINYEITNFALKGHESRHNKLNWELKPHLGLGLYAVSLLFCNDK 120
Db 300 SANFYRIASKRLSEAGYNHYEISSYCKPGYCKEHNLTYNQNRPFYAFGLGSASYI----- 354
Qy 121 NNVVR 125
Db 355 -NGVR 358

RESULT 13

US-10-289-762-396
; Sequence 396, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; CURRENT FILING DATE: 2003-03-27

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:21:03 ; Search time 84.2496 Seconds
(without alignments)
1966.133 Million cell updates/sec

Title: US-10-688-058-12

Perfect score: 1981

Sequence: 1 MRVDLPLVLSLYINLSFC.....LVKIIHFNDNPFVKVRLRP 377

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1981	100.0	377	8	ADU98753
2	1090	55.0	204	8	ADU98751
3	863	43.6	176	2	AAV20034
4	763	38.5	156	2	AAV20035
5	258.5	13.0	379	5	ABBS4455
6	255.5	12.9	408	5	ABP27526
7	250.5	12.6	376	6	ABU01854
8	250.5	12.6	376	8	ADK46982
9	249	12.6	376	5	ABP27525
10	248.5	12.5	409	8	ADR95677
11	248.5	12.5	409	9	AEA59547
12	244	12.3	376	8	ADV87963
13	244	12.3	376	8	ADV79216
14	244	12.3	376	8	ADV81406
15	237	12.0	382	9	AEBS3646
16	233.5	11.8	375	9	AEBA4038
17	232.5	11.7	385	6	ABM70978
18	231.5	11.7	385	6	ABJ18960
19	206.5	10.4	381	9	AEA49154
20	199	10.0	385	5	ABBA4803
21	196.5	9.9	391	5	ABP38845
22	196.5	9.9	393	7	ADC97174
23	190.5	9.6	397	7	ADH88000
24	185.5	9.4	391	6	ABP77997

25	185	9.3	377	6	ABM67734	Abm67734	Phototrhob
26	181	9.1	382	7	ADF04923	Adf04923	Bacterial
27	173	8.7	295	2	AY371169	Aay371169	Protein i
28	173	8.7	394	2	AAV34978	Aay34978	Chlamydia
29	152.5	7.7	387	7	ABO65431	Abob65431	Klebsiell
30	150.5	7.6	389	6	ADA36801	Ada36801	Acinetoba
31	144.5	7.3	442	4	AA933553	Aab93553	Human pro
32	144.5	7.3	442	8	ABM81907	Abm81907	Tumour-as
33	143.5	7.2	506	7	ABO74320	Abob74320	Pseudomon
34	142.5	7.2	401	9	ABM95708	Abm95708	M. xanthu
35	142	7.2	460	2	AAV35541	Aay35541	C. pneumo
36	140.5	7.1	424	7	ABO71949	Abob71949	Pseudomon
37	139.5	7.0	1182	3	AA182888	Aab18288	Plasmodiu
38	135.5	6.8	320	6	ABU19076	Abu19076	Protein i
39	135	6.8	320	2	AAV37405	Aay37405	Human pro
40	128.5	6.5	442	4	AA95029	Ab95029	Human pro
41	128	6.5	479	8	ADL04508	Adl04508	M. catarr
42	124.5	6.3	548	8	ADS08193	Ads08193	Staphyloc
43	124.5	6.3	605	6	ABU26712	Abu26712	Protein e
44	124	6.3	473	6	ABP80789	Abp80789	N. gonorr
45	123.5	6.2	4987	2	AAR10834	Aar10834	Rianodin

ALIGNMENTS

RESULT 1
ADU98753
ID ADU98753 standard; protein; 377 AA.
AC ADU98753;
XX
XX
DT 24-FEB-2005 (first entry)
DE Borrelia burgdorferi antigenic polypeptide seqid 12.
XX
XX
KW antibacterial; vaccine; immune stimulation; immunity; antigen;
KW DNA library.
XX
XX
OS Borrelia burgdorferi.
XX
XX
PN WQ2004103269-A2.
XX
XX
PD 02-DEC-2004.
XX
XX
PF 17-OCT-2003; 2003WO-US033056.
XX
XX
PR 18-OCT-2002; 2002US-0419401P.
XX
XX
PA (MACR-) MACROGENICS INC.
PA (TEXA) UNIV TEXAS SYSTEM.
XX
XX
PI Sykes KE, Hale KS, Johnston SA;
XX
XX
DR WPI; 2004-834155/82.
DR N-PSDB; ADU98752.
XX
XX
PT Immunizing a subject against Borrelia burgdorferi infection comprises
PT providing to the subject at least one Borrelia antigen or its fragment.
XX
XX
PS Claim 27; SEQ ID NO 12; 121pp; English.
XX
XX
CC The invention describes a method of immunizing a subject comprising
CC providing to the subject at least one Borrelia antigen or its fragment to
CC induce an immune response. Also described are: an isolated polynucleotide
CC comprising a sequence having at least 17 contiguous nucleotides in common
CC with a sequence not given in the specification; an isolated polypeptide
CC having at least 5 consecutive amino acids of the sequence not given in
CC the specification; a vaccine composition comprising at least one Borrelia
CC antigen or at least one polynucleotide encoding a Borrelia antigen;
CC screening for at least one test polypeptide or test polynucleotide
CC encoding a polypeptide for an ability to produce an immune response;
CC preparing a vaccine; vaccinating a subject; treating a subject infected

CC with a pathogen; passive immunization; vaccination; detecting Herpesvirus
CC ; and detecting antibodies to a herpesvirus. The method is useful in
CC immunizing a subject against Borrelia burgdorferi infection. This is the
CC amino acids sequence of a Borrelia burgdorferi polypeptide that may be
CC useful in immunizing a subject against infection by a member of the
CC Borrelia species.
XX
SQ Sequence 377 AA;
Query Match 100.0%; Score 1981; DB 8; Length 377;
Best Local Similarity 100.0%; Pred. No. 1.6e-169;
Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRVDLLPLVELSYINLSFCKDPSFNRLLEELKCHLILGHPIIKTLVKHVDCLSR 60
DB 1 MRVDLLPLVELSYINLSFCKDPSFNRLLEELKCHLILGHPIIKTLVKHVDCLSR 60
QY 61 QDNLKFTFTSLSKYINLELLEETLEIIPGVDFEFKLLDEFCTIRINLNVSFSLEFR 120
DB 61 QDNLKFTFTSLSKYINLELLEETLEIIPGVDFEFKLLDEFCTIRINLNVSFSLEFR 120
QY 121 KIVGIPISYKKNILINNIRKPPFDLNDMTVMNPLQKSHLKRDLKELLSYMPHEICF 180
DB 121 KIVGIPISYKKNILINNIRKPPFDLNDMTVMNPLQKSHLKRDLKELLSYMPHEICF 180
QY 181 SDFTCEEGFVLRDFOINSIDSEKLFCALECSNGYINYEITNFALKGHESRHNKLWNE 240
DB 181 SDFICEEGFVLRDFOINSIDSEKLFCALECSNGYINYEITNFALKGHESRHNKLWNE 240
QY 241 LKPHLGLGLYAVSLFLCNDKNNNRALIRKTGSFVKANNHLVTFELLEDLEFFVYHFIOG 300
DB 241 LKPHLGLGLYAVSLFLCNDKNNNRALIRKTGSFVKANNHLVTFELLEDLEFFVYHFIOG 300
QY 301 LGTIQGVSLRALRLRFYNEKQPFQFINYCSTLSKKFVFDNIMLLKGRERFKLNFYLVK 360
DB 301 LGTIQGVSLRALRLRFYNEKQPFQFINYCSTLSKKFVFDNIMLLKGRERFKLNFYLVK 360
QY 361 IINHFNDFPKVKLRLP 377
DB 361 IINHFNDFPKVKLRLP 377
RESULT 2
ADU98751
ID ADU98751 standard; protein; 204 AA.
AC ADU98751;
XX
XX 24-FEB-2005 (first entry)
DE Borrelia burgdorferi antigenic polypeptide seqid 10.
XX
XX antibacterial; vaccine; immune stimulation; immunity; antigen;
KW DNA library.
XX
XX Borrelia burgdorferi.
OS
XX
XX WO2004103269-A2.
XX
XX 02-DEC-2004.
XX
XX 17-OCT-2003; 2003WO-US033056.
XX
XX 18-OCT-2002; 2002US-0419401P.
XX
XX (MACR-) MACROGENICS INC.
XX (TEXA) UNIV TEXAS SYSTEM.
XX
XX Sykes KF, Hale KS, Johnston SA;
XX
XX WPI; 2004-834155/82.
XX
XX N-PSDB; ADU98750.
XX

PT Immunizing a subject against Borrelia burgdorferi infection comprises
PT providing to the subject at least one Borrelia antigen or its fragment.
XX
PS Claim 27; SEQ ID NO 10; 121pp; English.
XX
CC The invention describes a method of immunizing a subject comprising
CC providing to the subject at least one Borrelia antigen or its fragment to
CC induce an immune response. Also described are: an isolated polynucleotide
CC comprising a sequence having at least 17 contiguous nucleotides in common
CC with a sequence not given in the specification; an isolated polypeptide
CC having at least 5 consecutive amino acids of the sequence not given in
CC the specification; a vaccine composition comprising at least one Borrelia
CC antigen or at least one polynucleotide encoding a Borrelia antigen;
CC screening a polypeptide for an ability to produce an immune response;
CC encoding a vaccine; vaccinating a subject; treating a subject infected
CC with a pathogen; passive immunization; vaccination; detecting Herpesvirus
CC ; and detecting antibodies to a herpesvirus. The method is useful in
CC immunizing a subject against Borrelia burgdorferi infection. This is the
CC amino acids sequence of a Borrelia burgdorferi polypeptide that may be
CC useful in immunizing a subject against infection by a member of the
CC Borrelia species.
XX
SQ Sequence 204 AA;
Query Match 55.0%; Score 1090; DB 8; Length 204;
Best Local Similarity 100.0%; Pred. No. 1.2e-89;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 141 RKFPFDLNDMTVMNPLQKSHLKRDLKELLSYMPHEICFSDFICEEGFVLRDFOINSID 200
DB 1 RKFPFDLNDMTVMNPLQKSHLKRDLKELLSYMPHEICFSDFICEEGFVLRDFOINSID 60
QY 201 SEKLWFCALCELSNGYINYEITNFALKGHESRHNKLWELKPHLGLGLYAVSLFLCNDK 260
DB 61 SEKLWFCALCELSNGYINYEITNFALKGHESRHNKLWELKPHLGLGLYAVSLFLCNDK 120
QY 261 NNNVRALIRKTGSFVKANNHLVTFELLEDLEFFVYHFIOQLGTIQGVSLRALRLRFYNE 320
DB 121 NNNVRALIRKTGSFVKANNHLVTFELLEDLEFFVYHFIOQLGTIQGVSLRALRLRFYNE 180
QY 321 KQFFQFINYCSTLSKKFVFDNIM 344
DB 181 KQFFQFINYCSTLSKKFVFDNIM 204
RESULT 3
AAY20034
ID AAY20034 standard; protein; 176 AA.
XX
XX AAY20034;
AC
XX 19-JUL-1999 (first entry)
DT
XX
XX B. burgdorferi antigenic protein, f229.aa.
DE
XX
XX Antigenic protein; vaccine; Lyme disease; infection; detection.
KW Borrelia burgdorferi.
XX
XX OS
XX WO9859071-A1.
XX
XX 30-DEC-1998.
PD
XX
XX 18-JUN-1998; 98WO-US012718.
XX
XX 20-JUN-1997; 97US-0050359P.
PR 22-JUL-1997; 97US-0053344P.
XX
PR 22-JUL-1997; 97US-0053377P.
XX
PR 03-SEP-1997; 97US-0057483P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX (MEDI-) MEDIMMUNE INC.
PA

XX PI Choi GH, Erwin AL, Hanson MS, Lathigra R;
 XX DR WPI; 1999-189980/16.
 XX DR N-PSDB; AAX61731.
 XX PT New isolated Borrelia burgdorferi nucleic acids - used to develop
 XX PT products for the diagnosis, prevention and treatment of diseases caused
 XX PT by Borrelia, particularly Lyme disease.
 XX PS Claim 12; Page 169; 275pp; English.
 XX CC This sequence represents a Borrelia burgdorferi (Bb) protein of the
 CC invention, which is suitable for use in a vaccine. The Bb polypeptides
 CC can be used in vaccines for eliciting protective antibodies to members of
 CC the Borrelia genus, particularly for the use against Lyme disease in
 CC humans and animals. They can be used for preventing or attenuating an
 CC infection caused by a member of the Borrelia genus. The products can also
 CC be used for detection of members of the Borrelia genus
 XX SQ Sequence 176 AA;
 Query Match 43.6%; Score 863; DB 2; Length 176;
 Best Local Similarity 98.2%; Pred. No. 2.6e-69;
 Matches 167; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MRVLLPLVELSLYINLSFCCKDPSIFNRIIEELKCHLLILGHPIIKTYIKHVDVDFCLSR 60
 DB 1 MRVLLPLVELSLYINLSFCCKDPSIFNRIIEELKCHLLILGHPIIKTYIKHVDVDFCLSR 60
 QY 61 QDNLKFTSLSKYINLELLEEFTELEIIPGYVDPEKPKLLDEFCITRINLNVSFSLEFR 120
 DB 61 QDNLKFTSLSKYINLELLEEFTELEIIPGYVDPEKPKLLDEFCITRINLNVSFSLEFR 120
 QY 121 KIVGPEISYKMLNINIRKFFDLNIDMTVMPLQKSHLKRDLKEL 170
 DB 121 KIVGPEISYKMLNINIRKFFDLNIDMTVMPLQKSHLKRDLQRI 170
 RESULT 4
 AAY20035
 ID AAY20035 standard; protein; 156 AA.
 AC AAY20035;
 XX DT 19-JUL-1999 (first entry)
 XX DE B. burgdorferi antigenic protein, t229.aa.
 XX KW Antigenic protein; vaccine; Lyme disease; infection; detection.
 XX OS Borrelia burgdorferi.
 XX PN W09859071-A1.
 XX PD 30-DEC-1998.
 XX PF 18-JUN-1998; 98WO-US012718.
 XX PR 20-JUN-1997; 97US-0050359P.
 XX PR 22-JUL-1997; 97US-0053344P.
 XX PR 22-JUL-1997; 97US-0053377P.
 XX PR 03-SEP-1997; 97US-0057483P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PA (MEDI-) MEDIMMUNE INC.
 XX PI Choi GH, Erwin AL, Hanson MS, Lathigra R;
 XX DR WPI; 1999-189980/16.
 XX DR N-PSDB; AAX61732.
 XX PT New isolated Borrelia burgdorferi nucleic acids - used to develop

PT products for the diagnosis, prevention and treatment of diseases caused
 XX by Borrelia, particularly Lyme disease.
 XX PS Claim 12; Page 169; 275pp; English.
 XX CC This sequence represents a Borrelia burgdorferi (Bb) protein of the
 CC invention, which is suitable for use in a vaccine. The Bb polypeptides
 CC can be used in vaccines for eliciting protective antibodies to members of
 CC the Borrelia genus, particularly for the use against Lyme disease in
 CC humans and animals. They can be used for preventing or attenuating an
 CC infection caused by a member of the Borrelia genus. The products can also
 CC be used for detection of members of the Borrelia genus
 XX SQ Sequence 156 AA;
 Query Match 38.5%; Score 763; DB 2; Length 156;
 Best Local Similarity 98.0%; Pred. No. 2.2e-60;
 Matches 147; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 21 CKDPSIFNRIIEELKCHLLILGHPIIKTYIKHVDVDFCLSRQDNLKFTSLSKYINLELL 80
 DB 1 CKDPSIFNRIIEELKCHLLILGHPIIKTYIKHVDVDFCLSRQDNLKFTSLSKYINLELL 60
 QY 81 EETLEIIPGYVDPEKPKLLDEFCITRINLNVSFSLEFRKIVGIPISYKMLNINIRI 140
 DB 61 EETLEIIPGYVDPEKPKLLDEFCITRINLNVSFSLEFRKIVGIPISYKMLNINIRI 120
 QY 141 RKFPFDLNDMTVMPLQKSHLKRDLKEL 170
 DB 121 RKFPFDLNDMTVMPLQKSHLKRDLQRI 150
 RESULT 5
 ABB54455
 ID ABB54455 standard; protein; 379 AA.
 AC ABB54455;
 XX DT 29-AUG-2003 (revised)
 XX DT 16-MAY-2002 (first entry)
 XX DE Lactococcus lactis protein hemN.
 XX KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
 XX OS Lactococcus lactis; IL1403.
 XX PN FR2807446-A1.
 XX PD 12-OCT-2001.
 XX PF 11-APR-2000; 2000FR-00004630.
 XX PR 11-APR-2000; 2000FR-00004630.
 XX PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
 XX PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;
 XX DR WPI; 2002-043418/06.
 XX PT New nucleotide sequence useful in the identification of Lactococcus
 PT lactis and related species.
 XX PS Claim 6; SEQ ID NO 1157; 2504pp; French.
 XX CC The present invention is related to a Lactococcus lactis nucleotide
 CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The nucleic
 CC acid sequence is useful in the detection and/or amplification of nucleic
 CC acid sequence, particularly to identify Lactococcus lactis or related
 CC species. The proteins of the invention are useful for the biosynthesis or
 CC biodegradation of a composition of interest. The invention helps research
 CC in lactic bacteria, particularly useful in the production of yogurt and

Db 63 PTALSAPQLEVLNGLTNLDSVLEELTTEANPGDLADKIAVLKNSAVNRVSLGVQTF 122
Qy 116 SLEPRKIVGIPETSYKKNLILNIRKFPD-LNIDMTVNNPQKKSHLKRDLKELLSYM 174
Db 123 DDKWLKIKGRSHLB-KDIYENIDRLKLAGDNISIDLIYALPGQTMQOVKENVAKAIGLD 181
Qy 175 PEHICFSDFTICEERGFVLRDF-----DNSIDSEKLFWCALECLESNGYINYEITNFA 226
Db 182 IPHMSLSLILENHTVFNMRMRGKLPKPKEELEAEEMFEVITAE-LERAGFEHYEISNFS 240
Qy 227 LKGHESRHNKLWELKPHLGLGLYAVSLVLLFCNDKNNVR-----ALIRKTSFVKANHLV 282
Db 241 KPGFESRHNLMYDWAEBYYIGAGASGV-----NGVRYNKHGPIRHYLSAVEEGNACI 294
Qy 283 TFELLEDEFPVYHFIOGLGTIQGVSLRALRLRPEYNEKOFF 324
Db 295 TEDHLSQKEQMEEMFGLRKKSQVSM-----ARFEKFGQSF 332

RESULT 9
ABP27525
ID ABP27525 standard; protein; 376 AA.
XX AC ABP27525;
XX DT 02-JUL-2002 (first entry)
XX DE Streptococcus polypeptide SEQ ID NO 4226.
XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
XX KW group A streptococcus; Streptococcus pyogenes; antibacterial;
XX KW anti-inflammatory; infection; vaccine; meningitis; gene therapy.
XX OS Streptococcus agalactiae.
XX FN WO200234771-A2.
XX PD 02-MAY-2002.
XX PF 29-OCT-2001; 2001WO-GB004789.
XX PR 27-OCT-2000; 2000GB-00026333.
XX PR 24-NOV-2000; 2000GB-00028727.
XX PR 07-MAR-2001; 2001GB-00005640.
XX PA (CHIR-) CHIRON SPA.
XX PA (GENO-) INST GENOMIC RBS.
XX PI Telford J, Masignani V, Margarit Y Rosi, Grandi G, Fraser C;
XX PI Tettelin H;
XX XX
XX DR WPI; 2002-352536/38.
XX DR N-PSDB; ABN68156.
XX XX
XX PT New Streptococcus protein for the treatment or prevention of infection or
XX PT disease caused by Streptococcus bacteria, such as meningitis, and for
XX PT detecting a compound that binds to the protein.
XX XX
XX PS Claim 1; Page 3580; 4525pp; English.
XX XX
XX CC The invention relates to a protein (ABP25413-ABP30895) from group B
XX CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
XX CC (Streptococcus pyogenes), comprising one of 5493 sequences (S1), given in
XX CC the specification. The proteins have antibacterial and anti-inflammatory
XX CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
XX CC antibodies that bind (I) are used in the manufacture of medicaments for
XX CC the treatment or prevention of infection or disease caused by
XX CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
XX CC Nucleic acids encoding (I) are used to detect Streptococcus in a
XX CC biological sample. (I) is used to determine whether a compound binds to
XX CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
XX CC used as a vaccine or diagnostic composition. The disease caused by

CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX Sequence 376 AA;
SQ Query Match 12.6%; Score 249; DB 5; Length 376;
Best Local Similarity 24.9%; Pred. No. 1.3e-13;
Matches 87; Conservative 68; Mismatches 145; Indels 50; Gaps 12;
Qy 12 SLVINUSPCK-----DFS-----IFNRILEELKCHLILGHPIKTIYIKHVD 55
Db 7 SAYVHIPFCTQICYYCDFSKVFKNQPVDAVLAQALIREFRSYDI-----TELRLTYIGGGT 62
Qy 56 FCLSRQDNLFKFTSLSKYINLELLEEFTELEITPQVDFEKKPLDDEFCTIRNLNVQSF 115
Db 63 PTSISAVQDYLTLTELSRDNLNLTLEEFTEANPGDLTVDKIEVLQKSAVNRVSLGVQTF 122
Qy 116 SLEFRKIVGIPETSYKKNLILNIRKFPF-DLNIDMTVNNPQKKSHLKRDLKELLS 172
Db 123 NDKHLKRIG---RSHNEAQIYSTIDALKTAGFQNISIDLIYALPGQTMDDVRSNVAKALS 179
Qy 173 YMPEHICFSDFTICEERGFVLRDF-----DNSIDSEKLFWCALECLESNGYINYEITN 224
Db 180 LNIPLHLSYLSLILEHHTVFNMRMRGKLPKPKEELEAEEMFEVITAE-LERAGFEHYEISN 238
Qy 225 FALKGHESRHNKLWELKPHLGLGLYAVSLVLLFCNDKNNVRALIRKTSF-----VKAN 278
Db 239 FTKPGFESRHNLMYDWAEBYYIGAGASGVL-----DGIR--YRNRGPIQHYLKGVSSEG 290
Qy 279 NHLVTFELLEDEFPVYHFIOGLGTIQGVSLRALRLRFEYN-EKQFPQFI 327
Db 291 NARLSEVLKSNEMMEELFLGLRKKEGVSGIKFEQKFGTSFEKRYQIV 340

RESULT 10
ADR95677
ID ADR95677 standard; protein; 409 AA.
XX AC ADR95677;
XX DT 16-DEC-2004 (first entry)
XX DE Novel S. pneumoniae protein sequence, SEQ ID 4312.
XX KW Meningitis; bacteraemia; pneumonia; otitis media; vaccine;
XX KW bacterial infection.
XX OS Streptococcus pneumoniae.
XX XX
XX EN US6800744-B1.
XX PD 05-OCT-2004.
XX PF 30-JUN-1998; 98US-00107433.
XX PR 02-JUL-1997; 97US-0051553P.
XX PR 12-MAY-1998; 98US-0085131P.
XX XX
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX XX
XX PI Doucette-Stamm LA, Bueh D;
XX XX
XX DR WPI; 2004-697205/68.
XX DR N-PSDB; ADR93074.
XX XX
XX PT New isolated nucleic acid encoding a Streptococcus pneumoniae
XX PT polypeptide, useful for diagnosing, preventing and/or treating
XX PT pathological conditions resulting from the bacterial infection.
XX PS Disclosure; SEQ ID NO 4312; 151pp; English.

XX The invention relates to an isolated nucleic acid comprising a sequence
CC encoding a Streptococcus pneumoniae ADR91366polypeptide, or its
CC fragments, with any of 9 fully defined sequences (appearing as ADR94308,
CC ADR94489, ADR94800, ADR94837, ADR94969, ADR95253, ADR95642, ADR95682,
CC ADR96079) or any of the fully defined sequences appearing as ADR91705,
CC ADR91886, ADR92197, ADR92234, ADR93039, ADR93079, ADR92366, ADR92850 or
CC ADR93476 or at least 20 or 30 consecutive nucleotides of the nucleotide
CC sequences, or at least 40, 60 or 300 consecutive nucleotides, which is
CC hybridisable under high stringency conditions to the nucleotide sequences.
CC The nucleic acids and proteins are chosen from 5206 disclosed sequences.
CC Also included are a recombinant expression vector comprising the isolated
CC nucleic acid cited above operably linked to a transcription regulatory
CC element, a cell comprising the recombinant expression vector and a probe
CC comprising at least 20 consecutive nucleotides of the nucleotide
CC sequences as cited above. The methods and compositions of the present
CC invention are useful for the diagnosis, prevention and/or treatment of
CC pathological conditions resulting from bacterial infection by
CC Streptococcus pneumoniae e.g. pneumonia, bacteraemia, meningitis and
CC otitis media. The present sequence is one of the 2603 disclosed S.
CC pneumoniae protein sequences. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=68007481.
XX

Query Match 12.5%; Score 248.5; DB 8; Length 409;
Best Local Similarity 25.7%; Pred. No. 1.6e-13;
Matches 88; Conservative 58; Mismatches 151; Indels 45; Gaps 10;
QY 12 SLYNLSFCCCK-----DF-----SIFNRILEELKCHLLGHPIIKTLYIKHVD 55
DB 40 SAYVHIPFCTQICVYCDFSKVFKNQPVDSVLEHLLLEFRSYDI-----EKLSTLYIRGWN 95
QY 56 FCLSRQDNLFKIFTSLSKYINLELLEBFTLEIIPGYVDFFKFLDLBFCITRNLNVQSF 115
DB 96 TTALSAPQLDVLNGLTINLDLSVLEELTEANPGDLADKIAVLKNSAVNRVSLGVQTF 155
QY 116 SLEFRKIVGPIEISYKKNILNINIRKFPDP-LNIDMTVMPLQKSHLKDLELLSYM 174
DB 156 DDKMLKKIGRSHLE-KDIYENIDRLKLAGFDNISIDLIYALPGQTMQVKENVAKGLD 214
QY 175 PEHICFSDIFCEESEGFVLDRP-----DNSIDSEKLFWCALECELESGYINYEITNEA 226
DB 215 IPHMSLSLILENHTVFMNRRGKPLPKEELEAMFEYIIAE-LERAGFEHYEISNFS 273
QY 227 LKGHESRHNKLNWELKPHLGLGLYAVSLFCNDKNNVR-----ALIRKTSFVKANNHLV 282
DB 274 KPGFESRHNLMYDWAEEYIGACASGVY-----NGVRYKNGPIRHYLSAVEEGNACI 327
QY 283 TFELLEDEFFVYHFIQSLGTIQGVSLRALRLFEYNEKQPF 324
DB 328 TEDHLSQKEQMEEMFLGRKKSGVM-----ARFEKFGQSF 365

RESULT 11
ID AEA59547
XX AEA59547 standard; protein; 409 AA.
XX AEA59547;
XX 25-AUG-2005 (first entry)
XX Streptococcus pneumoniae ORF amino acid sequence SEQ ID NO:4312.
XX bacterial infection; Streptococcus pneumoniae infection; antibacterial;
XX vaccine.
XX Streptococcus pneumoniae.
XX US2005136404-A1.
XX

PD 23-JUN-2005.
XX 10-JUL-2003; 2003US-00617320.
XX 02-JUL-1997; 97US-0051553P.
PR 12-MAY-1998; 98US-0085131P.
PR 30-JUN-1998; 98US-00107433.
XX (DOUC/) DOUCETTE-STAMM L A.
PA (BUSH/) BUSH D.
XX Doucette-Stamm LA, Bush D;
XX WPI; 2005-477576/48.
DR N-PSDB; AEA56944.
XX New isolated nucleic acid molecules and encoded polypeptides useful for
PT diagnosing, preventing or treating bacterial infections, particularly
PT Streptococcus pneumoniae infection.
XX Claim 5; SEQ ID NO 4312; 144pp; English.
XX The invention relates to an isolated nucleic acid molecule for detecting,
CC preventing or treating pathological conditions resulting from bacterial
CC infection. The isolated nucleic acid comprises: (a) any of the 2603
CC nucleotide sequences of AEA55236 to AEA57838; (b) a nucleotide sequence
CC encoding a Streptococcus pneumoniae polypeptide comprising any of the
CC 2603 amino acid sequences of AEA57839 to AEA60441; or (c) a nucleotide
CC sequence of at least 8 nucleotides in length, where the sequence is
CC hybridisable to a nucleic acid having any of the nucleotide sequences in
CC (a). Also described: (1) a recombinant expression vector comprising the
CC above nucleic acid operably linked to a transcription regulatory element;
CC (2) a cell comprising the recombinant expression vector; (3) producing an
CC S. pneumoniae polypeptide; (4) a probe comprising a nucleotide sequence
CC consisting of at least 8 nucleotides of any of AEA55236 to AEA57838; (5)
CC treating a subject for S. pneumoniae infection; (6) a recombinant or
CC substantially pure preparation of an S. pneumoniae polypeptide or its
CC fragment, where the polypeptide is selected from AEA57839 to AEA60441;
CC (7) a vaccine composition for preventing or treating an S. pneumoniae
CC infection, comprising an amount of the above nucleic acid or polypeptide;
CC (8) detecting the presence of a Streptococcus nucleic acid in a sample;
CC (9) a computer readable medium having recorded the nucleotide sequences
CC of AEA55236 to AEA57838; (10) a computer based system for identifying
CC fragments of the Streptococcus genome of commercial importance. The
CC composition and methods are useful for diagnosing, preventing or treating
CC bacterial infections, particularly S. pneumoniae infection. The present
CC sequence represents a S. pneumoniae ORF amino acid sequence from the
CC present invention. Note - The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from the USPTO web site.
XX

Sequence 409 AA;
Query Match 12.5%; Score 248.5; DB 9; Length 409;
Best Local Similarity 25.7%; Pred. No. 1.6e-13;
Matches 88; Conservative 58; Mismatches 151; Indels 45; Gaps 10;
QY 12 SLYNLSFCCCK-----DF-----SIFNRILEELKCHLLGHPIIKTLYIKHVD 55
DB 40 SAYVHIPFCTQICVYCDFSKVFKNQPVDSVLEHLLLEFRSYDI-----EKLSTLYIRGWN 95
QY 56 FCLSRQDNLFKIFTSLSKYINLELLEBFTLEIIPGYVDFFKFLDLBFCITRNLNVQSF 115
DB 96 TTALSAPQLDVLNGLTINLDLSVLEELTEANPGDLADKIAVLKNSAVNRVSLGVQTF 155
QY 116 SLEFRKIVGPIEISYKKNILNINIRKFPDP-LNIDMTVMPLQKSHLKDLELLSYM 174
DB 156 DDKMLKKIGRSHLE-KDIYENIDRLKLAGFDNISIDLIYALPGQTMQVKENVAKGLD 214
QY 175 PEHICFSDIFCEESEGFVLDRP-----DNSIDSEKLFWCALECELESGYINYEITNEA 226
DB 215 IPHMSLSLILENHTVFMNRRGKPLPKEELEAMFEYIIAE-LERAGFEHYEISNFS 273


```

CC contains only 2344 sequences.
XX
SQ Sequence 376 AA;

Query Match      12.3%; Score 244; DB 8; Length 376;
Best Local Similarity 24.6%; Pred. No. 3.6e-13;
Matches 86; Conservative 68; Mismatches 146; Indels 50; Gaps 12;

QY 12 SLVINSFCK-----DFS-----IFNRILEELKCHLILGHPIIKTYIKHVD 55
DQ 7 SAYVHIPFCTQICYYCDFSKVFNQPDVAYLQALIREFRSYDI---TELRTLYIGGGT 62
QY 56 FCLSRQDNLFKFTSLSKVINLELLEEFTELEIIPGVYDPEKFLDLDEFCTRINLVQSF 115
DQ 63 PTSISAVQLDYLLTELIRDNLNLTLEEFTEIANPGLTVDKIEVLQKSAVNRVSLGVQTF 122
QY 116 SLERFKVIGPEISYKKNL--LNNIRKFPF--DLNIDMTVNNPQKKSHKRDKLKLLS 172
DQ 123 NDKHLKRIG---RSHNEAQIYSTIDALKTAGFQNISIDLIYALPGQTMDDVRNVAKALS 179
QY 173 YMEPHICFSDPICEEGFVLDRF-----DNSIDSEKLMWFCALCLESNGYINYEITN 224
DQ 180 LNIPHLSLYSLILEHHTVFNKMRGKHLPTEDLEAEMFEYIISE--MERNGFEHVEISN 238
QY 225 FALKGHESRHNKLNWELKPHGLGLYAVSLLFCNDKNNVRLIRKTGSP-----VKAN 278
DQ 239 FTKPGFESRNLNWDNVVEYGVGAGAGYL-----DGIR--YRNRGPIQHYLKGVSEG 290
QY 279 NHLVTFELLEDLEFFVYHFHQGLGTIQGVSLRALRLRFEYN--EKQFFQFI 327
DQ 291 NARLSEVLSKNEMWEEELFLGRKKEGVSIGKFEQKFGTSFKRYGQIV 340

RESULT 14
ADV81406
ID ADV81406 standard; protein; 376 AA.
XX
AC ADV81406;
XX
DT 24-FEB-2005 (first entry)
XX
DE Streptococcus agalactiae protein, SEQ ID 2547.
XX
KW Antibacterial; vaccine; bacterial infection.
XX
OS Streptococcus agalactiae.
XX
PN WO200292818-A2.
XX
PD 21-NOV-2002.
XX
PF 26-APR-2002; 2002WO-IB003059.
XX
PR 26-APR-2001; 2001FR-00005642.
XX
PA (INSP ) INST PASTEUR.
XX
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI Glaser P, Rusniok C, Chevalier F, Frangeul L, Lalioui L;
PI Zouine M, Couve E, Buchrieser C, Poyart C, Trieu-Cuot P, Kunst P;
XX
DR WPI; 2004-101891/11.
XX
XX Genomic nucleotide sequences encoding polypeptides of Streptococcus
PT agalactiae for the development of vaccines, diagnostic tools, DNA chips
PT and identification of therapeutic targets.
XX
PS Claim 6; SEQ ID NO 2547; 439pp; French.
XX
CC The present invention relates to novel Streptococcus agalactiae
CC nucleotide sequences (I; ADV78860-ADV78998 and ADV83341-ADV85476) and
CC novel polypeptides (II; ADV78999-ADV81203 and ADV81205-ADV83340). The
CC nucleotide sequences encode polypeptides of S. agalactiae involved in the

```

```

CC synthesis of amino acids, cell membranes, intermediate (central)
CC metabolism, energetic metabolism, fatty acid and phospholipid metabolism,
CC nucleotide metabolism including purines, pyrimidines and/or nucleosides,
CC regulatory functions, replication, transcription, translation, protein
CC transport, adaptation to atypical conditions, sensitivity to medicines
CC and/or analogues, functions related to transporters, biosynthesis of
CC cofactors, prosthetic groups and transporters, cell membrane proteins and
CC cellular machinery. (I) are useful for the detection and/or amplification
CC of nucleic acids. Pharmaceutical composition comprising (I) or (II) are
CC useful for treatment of a bacterial S. agalactiae infection. The complete
CC genome of Streptococcus agalactiae is given in ADV81204. Note: The
CC present patent is an equivalent for the basic patent FR2824074A1, which
XX contains only 2344 sequences.
SQ Sequence 376 AA;

Query Match      12.3%; Score 244; DB 8; Length 376;
Best Local Similarity 24.6%; Pred. No. 3.6e-13;
Matches 86; Conservative 68; Mismatches 146; Indels 50; Gaps 12;

QY 12 SLVINSFCK-----DFS-----IFNRILEELKCHLILGHPIIKTYIKHVD 55
DQ 7 SAYVHIPFCTQICYYCDFSKVFNQPDVAYLQALIREFRSYDI---TELRTLYIGGGT 62
QY 56 FCLSRQDNLFKFTSLSKVINLELLEEFTELEIIPGVYDPEKFLDLDEFCTRINLVQSF 115
DQ 63 PTSISAVQLDYLLTELIRDNLNLTLEEFTEIANPGLTVDKIEVLQKSAVNRVSLGVQTF 122
QY 116 SLERFKVIGPEISYKKNL--LNNIRKFPF--DLNIDMTVNNPQKKSHKRDKLKLLS 172
DQ 123 NDKHLKRIG---RSHNEAQIYSTIDALKTAGFQNISIDLIYALPGQTMDDVRNVAKALS 179
QY 173 YMEPHICFSDPICEEGFVLDRF-----DNSIDSEKLMWFCALCLESNGYINYEITN 224
DQ 180 LNIPHLSLYSLILEHHTVFNKMRGKHLPTEDLEAEMFEYIISE--MERNGFEHVEISN 238
QY 225 FALKGHESRHNKLNWELKPHGLGLYAVSLLFCNDKNNVRLIRKTGSP-----VKAN 278
DQ 239 FTKPGFESRNLNWDNVVEYGVGAGAGYL-----DGIR--YRNRGPIQHYLKGVSEG 290
QY 279 NHLVTFELLEDLEFFVYHFHQGLGTIQGVSLRALRLRFEYN--EKQFFQFI 327
DQ 291 NARLSEVLSKNEMWEEELFLGRKKEGVSIGKFEQKFGTSFKRYGQIV 340

RESULT 15
AEB36646
ID AEB36646 standard; protein; 382 AA.
XX
AC AEB36646;
XX
DT 08-SEP-2005 (first entry)
XX
DE L. pneumophila protein SEQ ID NO 978.
XX
KW detection; infection; Antibacterial; Vaccine.
XX
OS Legionella pneumophila.
XX
PN WO2005049642-A2.
XX
PD 02-JUN-2005.
XX
PF 23-SEP-2004; 2004WO-IB003578.
XX
PR 21-NOV-2003; 2003FR-00013687.
XX
PA (INSP ) INST PASTEUR.
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA (UPLY-) UNIV LYON 1 BERNARD CLAUDE.
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI Buchrieser C, Tichit M, Etienne J, Ma L, Cazalet C, Glaser P;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:28:54 ; Search time 14.3023 Seconds
(without alignments)
2536.214 Million cell updates/sec

Title: US-10-688-058-12
Perfect score: 1981
Sequence: 1 MRVDLLPLVELSLYINLSFC.....LVKLIINHFNDFPKVKLRLP 377

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: Pirl:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	100.0	377	2	G70181	oxygen-independent
2	258.5	13.0	379	C86767	oxygen-independent
3	253	12.8	396	AD2425	coproporphyrinogen
4	251	12.7	374	G97057	coproporphyrinogen
5	250.5	12.6	376	A98030	coproporphyrinogen
6	250.5	12.6	376	B95164	hypothetical prote
7	249	12.6	385	G83817	coproporphyrinogen
8	236	11.9	374	D72288	hypothetical prote
9	232	11.7	374	F89939	hypothetical prote
10	222.5	11.2	370	H70423	oxygen-independent
11	202.5	10.2	366	B69640	coproporphyrinogen
12	201.5	10.2	393	D71354	probable oxygen-in
13	199	10.0	385	AD1259	coproporphyrinogen
14	199	10.0	385	AH1621	coproporphyrinogen
15	193	9.7	399	C71728	probable oxygen-in
16	190.5	9.6	376	AB0116	probable oxygen-in
17	190	9.6	400	C97727	hypothetical prote
18	189	9.5	378	B71564	probable porphyrin
19	184.5	9.3	399	C81932	probable porphyrin
20	181.5	9.2	376	B84934	hypothetical prote
21	180.5	9.1	354	F82911	oxygen-independent
22	180.5	9.1	391	B81172	oxygen-independent
23	179.5	9.1	391	G82320	probable oxygen-in
24	178	9.0	383	B64070	coproporphyrinogen
25	169.5	8.6	412	S75358	coproporphyrinogen
26	169	8.5	378	AH0878	probable oxygen-in
27	167.5	8.5	378	G91107	probable oxidase {
28	166.5	8.4	373	B86538	coproporphyrinogen
29	166.5	8.4	373	A72086	coproporphyrinogen

30	166.5	8.4	378	2	B65081	hypothetical prote
31	165	8.3	345	2	E71843	oxygen-independent
32	164	8.3	352	1	B64673	coproporphyrinogen
33	164	8.3	383	2	E82673	conserved hypothet
34	163.5	8.3	378	2	B85953	probable oxidase y
35	156	7.9	457	2	E81739	oxygen-independent
36	155.5	7.8	388	2	B75557	probable coproporp
37	152	7.7	355	2	C81405	probable oxidoredu
38	149	7.5	457	2	B71476	probable coproporp
39	148.5	7.5	399	2	AE2616	hypothetical prote
40	148.5	7.5	402	2	E97398	probable oxygen-in
41	145.5	7.3	949	2	F90086	chromosomal region
42	143.5	7.2	460	2	A83454	oxygen-independent
43	142	7.2	458	2	E72023	oxygen-independent
44	142	7.2	458	2	G86601	coproporphyrinogen
45	139.5	7.0	384	2	A83597	probable oxidase P

ALIGNMENTS

RESULT 1

G70181

oxygen-independent coproporphyrinogen III oxidase homolog - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C;Accession: G70181

R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997

A;Authors: Smith, H.O.; Venter, J.C.

A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A;Reference number: A70100; MUID:98065943; PMID:9403685

A;Accession: G70181

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-377 <KLE>

A;Cross-references: UNIPROT:O51600; UNIPARC:UPI0000057576; GB:AE001167; GB:AE000783; NID

A;Experimental source: strain B31

Query Match	100.0%	Score	1981;	DB	2;	Length	377;
Best Local Similarity	100.0%	Pred. No.	1.4e-128;				
Matches	377;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MRVDLLPLVELSLYINLSFCCKDFSI	FNRLIEELKCHLILGHPIIKTLVIKHVDFCLSR	60			
Db	1	MRVDLLPLVELSLYINLSFCCKDFSI	FNRLIEELKCHLILGHPIIKTLVIKHVDFCLSR	60			
Qy	61	QDNLKFTFTSLSKYINLELEEFTE	IIPGYVDFEKFPLDDEFCTITRINLNVSFSLEFR	120			
Db	61	QDNLKFTFTSLSKYINLELEEFTE	IIPGYVDFEKFPLDDEFCTITRINLNVSFSLEFR	120			
Qy	121	KIVGIPETISKYKLNILINNIRKPP	FDLNDTMTNMPLOKKSHLKRDLKELLSYMPHEICF	180			
Db	121	KIVGIPETISKYKLNILINNIRKPP	FDLNDTMTNMPLOKKSHLKRDLKELLSYMPHEICF	180			
Qy	181	SDPFCSEEGVLRDFDONSIDSEKL	WFCALCESNGYINYEITNEALKGHESHKLNWE	240			
Db	181	SDPFCSEEGVLRDFDONSIDSEKL	WFCALCESNGYINYEITNEALKGHESHKLNWE	240			
Qy	241	LKPHLGLGLYAVSLLFCNDKNNV	RALIRKTSGFVKANNHLVTFELLEDFEYVHFIOG	300			
Db	241	LKPHLGLGLYAVSLLFCNDKNNV	RALIRKTSGFVKANNHLVTFELLEDFEYVHFIOG	300			
Qy	301	LGTIQGVSLRALRREYNEKQFFQ	YINYSTLSKPFVFDNIMLLKGRFRKLNFLYVK	360			
Db	301	LGTIQGVSLRALRREYNEKQFFQ	YINYSTLSKPFVFDNIMLLKGRFRKLNFLYVK	360			
Qy	361	IINHFNDFPKVKLRLP	377				
Db	361	IINHFNDFPKVKLRLP	377				

RESULT 2

C86767
oxygen-independent coproporphyrinogen III oxidase [imported] - Lactococcus lactis subsp. 1
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: C86767
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: C86767
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-379 <STO>
A;Cross-references: UNIPROT:Q9CGF7; UNIPARC:UPI000000C69B3; GB:AE005176; PID:gl2724101; F
A;Experimental source: strain 11L403
C;Genetics:
A;Gene: hemN
C;Superfamily: oxygen-independent coproporphyrinogen oxidase

Query Match 13.0%; Score 258.5; DB 2; Length 379;
Best Local Similarity 22.8%; Pred. No. 1.6e-10;
Matches 91; Conservative 78; Mismatches 133; Indels 97; Gaps 15;
Qy 12 SLXINLSFCK-DFSIFNRILEELKCHLILGHPI-----IKT 48
Db 7 SAYPHIFCSHCYYCDFA-----KVLMTGQPIDAYTESLIEEFSQSPFIEKRT 55
Qy 49 LYIKHVPFCLSRQDNLFIFTSLSKYINLELEFTEIIPGYVDFPKFKLLDFCITRI 108
Db 56 IYIGGGTPSVLSAQQLERLLTAIAEQLDLEVLVEFTVEANPGDLSDEVIKVLADSAVNRI 115
Qy 109 NLNVQSFLEPRKIVGIPETISYKKNILINIRKFPD-LNIDMTVMNPLOKSHLKRDL 167
Db 116 SLGVQTFNNALKKIGRTHTEVQVYD-SVERLKKAGFENITIDLIYALPGQTMEMVSDV 174
Qy 168 KELLSYMPEHICFSDICEE-----EGFV-LRQFDNSIDSEKLFWCALECLENGY 217
Db 175 EKFLLELPHVALVSLILEDTVMNQRRGLLPLSPEDKNAD---MYEIMDLAKNGY 231
Qy 218 INYEITNPFALKGHESRNKLNWELKPHLGLGLYAVSLFCNDKNNVRALIRKTGSFVKA 277
Db 232 NHYEVSFGLPGFESKENITWDNEEYVGIGAGAGYL-----AGIR----- 273
Qy 278 NNHLVTELEDDLEFFVHFQIGLTIQGVSLRALRLRFEYNEKQFFQFNYCSTLSKKF 337
Db 274 -----YKNLGP---VHHYLKAAPTKRIINEEVLKSKSQTEEMFL-----GLRKK- 315
Qy 338 VFDDNIMLLKGRERFKLNF-----YLVKLIINH---FND 367
Db 316 ---SGVLVEKFPENPKCSFEKLYGEQITELINQKLLYND 351

RESULT 3

AD2425
coproporphyrinogen III oxidase [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AD2425
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shampo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AD2425
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-396 <KUR>
A;Cross-references: UNIPROT:Q8YMH7; UNIPARC:UPI000000CECD; GB:BA0000019; PIDN:BA076655.1;
A;Experimental source: strain PCC 7120
C;Genetics:

A;Gene: all4956

C;Superfamily: oxygen-independent coproporphyrinogen oxidase

Query Match 12.8%; Score 253; DB 2; Length 396;
Best Local Similarity 25.2%; Pred. No. 4e-10;
Matches 91; Conservative 65; Mismatches 159; Indels 46; Gaps 13;
Qy 6 LPVELSLVINLSFCK-DFSIF--NRILEE-----LKCHLILL---GH 43
Db 6 VPSLASSAVHIIPFCRRRCFYCDPIFVVDGDRGRTSVTISGYVDVLCEIAITPAFGQ 65

Qy 44 PIITLYIKHVPCLSRQDNLFIFTSLSKYINLELEFTEIIPGYVDFPKFKLLDEF 103
Db 66 P-LKTVFFGGGTPSLLSTEQAQLVTLQKFGIAPDAEISMEVDGTFDLAHIQGYSV 124
Qy 104 CITRINLVQSFLEPRKIVGIPETISYKKNILINIRKFPF-DLNDMTVMNPLOKSH 162
Db 125 GNVRSVLGVQAFQEBELLKVAGRSH-SLKQIFAADLIHQVEIPFESIDLSGLPHQSLDQ 183
Qy 163 LKRDLLKELLSYMPHEHICFSDICEE--GFVLRDFDNSIDSE---KMFCALECLESN 215
Db 184 WQSLDTAVNIAPTHISYIDLTIETPGTAFGRYKPGDNPLPTDETTVMYQMGOKILTGG 243
Qy 216 GYINYEITNPFALKGHESRNKLNWELKPHLGLGLYAVSLFCNDKNNVRALIRKTGSFV 275
Db 244 DYEHYEISYNAKPGHQCRHNRVYWNRPYYGFGMGAASYV---EGKRFTRP--RKTKEYY 298
Qy 276 K-----NNHLVTELEDDLEFFVHFQIGLTIQGVSLRALRLRFEYNEKQFFQFNY 329
Db 299 QWYQELIANHGVIDWEITPKADVLLLETMLGLADGVSLAALTEEF---GKEKIOELHQ 355
Qy 330 C 330
Db 356 C 356

RESULT 4

G97057
coproporphyrinogen III oxidase [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: G97057
R;Nolling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: G97057
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-374 <KUR>
A;Cross-references: UNIPROT:Q97JJ5; UNIPARC:UPI000000CAL57; GB:AE001437; PIDN:AAK79250.1;
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC1279
C;Superfamily: oxygen-independent coproporphyrinogen oxidase

Query Match 12.7%; Score 251; DB 2; Length 374;
Best Local Similarity 26.5%; Pred. No. 5.2e-10;
Matches 93; Conservative 62; Mismatches 142; Indels 54; Gaps 15;
Qy 10 ELSIYINLSFC-----CKDPSIFNRIELKCHLILGHPIIKTLVIKHVDFCLSRQDN 63
Db 3 KGLYIHIHPCKSKLYC-DFPSYACIED----FYI-----PYAKALCTETEKASINKVFS 53

Qy 64 LKFTFTSLSYKINLELE-----EFTLEIIPGYVDFPKFKLLDFCITRI 108
Db 54 SIFIGGGTPSFLSIEALNILCDALKVKRKTKDVEFTVEGNPNSFTEKLMWFKDMGVNRL 113
Qy 109 NLNVQSFLEPRKIVGIPETISYKKNILINIRKFPF-DLNDMTVMNPLOKSHLKRDL 167
Db 114 SIGLQACQDRLKLGRIH-TLKEFTVAFKRNRLGNFNINVDLMFGIPDQTLDFEKSL 172

Qy 168 KELLSYNPEHICFSDFTICE-----EGFVLDRDFNSIDSEKLMFCALECSNGYIN 219
Db 173 EFITKLPKPEHISYSLIVEGTPYFQKNEGKLGK-LPNEDEERDMYSFARTFLEEKGYNQ 231
Qy 220 YEITNPAKGHESRHNKLNELKPHIGLGLYAVSLLEFCND---KN-NVNRALIRK--TGS 273
Db 232 YEISNFAVKDKECHNLIYWELDNYIGCGASAHG--YFNGVRYRNINNVNKKYIEQISKGN 289
Qy 274 FVKANNHLVTPELLEDELEFFVYFIQGLGTIQGVSLRALRLRFEYNEKQPF 324
Db 290 SVVEENHRL--LKEDMEEFMF---LGLRKTGVSIEEFKLFKFKNDIQEVY 335

RESULT 5
A98030
C:coporophyrinogen oxidase (EC 1.3.3.3) [imported] - Streptococcus pneumoniae (strain R6
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: A98030
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; UID:21429245; PMID:11544234
A:Accession: A98030
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-376 <KUR>
A:Cross-references: UNIPROT:Q97Q22; UNIPARC:UPI0000051830; GB:AE007317;
C:Genetics:
A:Gene: hemN
C:Superfamily: oxygen-independent coporophyrinogen oxidase
C:Keywords: oxidoreductase

Query Match 12.6%; Score 250.5; DB 2; Length 376;
Best Local Similarity 26.0%; Pred. No. 5.6e-10;
Matches 89; Conservative 57; Mismatches 151; Indels 45; Gaps 10;
Qy 12 SLYNLSFCCK-----DF-----SIFNRILSELKCHLILGHPIIKTYIKHVD 55
Db 7 SAYVHIPFCTQICYCDFSKVFIKNQPVDSYLEHLLSEFRSYDI-----EKLSTLYIGGGT 62
Qy 56 FCLSRQDNLFKFTSLSKYINLELLEFTLEIIPGYVDPEKFLDLDEFCTIRINLVQSF 115
Db 63 PTALSAPOLEVLNGLTKNLDLSVLELTIEANPGDLADKIAVLKNSAVNRVSLGVQTF 122
Qy 116 SLEFRKIVGPIEISYKKNILNINIRKPPDP-LNIDMTVNMPLOKSHLKDRLKELLSYM 174
Db 123 DDKMLKKIGRSHLE-KDIYENIDRLKLAGFNSIDLIYALPGQTMEQVKENAKAIGLD 181
Qy 175 PEHICFSDICEEGFVLDRF-----DNSIDSEKLMFCALECSNGYINVEITNEA 226
Db 182 IPHMSLSLIDENHTVFNMRMRGKLPKPEEAEEMFEYIIAE-LERAGFEHYEISNFS 240
Qy 227 LKGHESRHNKLNELKPHIGLGLYAVSLLEFCNDKNNNR-----ALIRKTSFVKANNHLV 282
Db 241 KPGFESRHNLMYDNEAYYIGAGASGV-----NGVRYKHGPIRHVLSAVEEGNACI 294
Qy 283 TFELLEDELFVYHFYFIQGLGTIQGVSLRALRLRFEYNEKQPF 324
Db 295 TEDHLSQKEQMEEMFGLRKKSGVSM---ARFEKFGQSF 332

RESULT 7
G83817
C:coporophyrinogen iii oxidase hemN [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: G83817
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; UID:20512582; PMID:11058132
A:Accession: G83817
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-385 <STO>
A:Cross-references: UNIPROT:Q9KD75; UNIPARC:UPI00000C3B7C; GB:AF001511; GB:BA000004; NID
A:Experimental source: strain C-125
C:Genetics:
A:Gene: hemN
C:Superfamily: oxygen-independent coporophyrinogen oxidase

Query Match 12.6%; Score 249; DB 2; Length 385;
Best Local Similarity 25.5%; Pred. No. 7.3e-10;
Matches 100; Conservative 60; Mismatches 156; Indels 76; Gaps 17;
Qy 14 YINLSFC-----CKDPSIF-----NRILEELKCH-LILLGHPI--IKTYIKHVD 56
Db 6 YIHIPFCEHCYCYC-DPNKRYLKNQPNVEYLQALETEMAMVVAEQPKYSQTLVGGGTP 64
Qy 57 CLSRQDNLFKFTSLSKYINLELLEFTLEIIPGYVDPEKFLDLDEFCTIRINLVQSF 116

nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; UID:21357209; PMID:11463916
A:Accession: B95164
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-376 <KUR>
A:Cross-references: UNIPROT:Q97Q22; UNIPARC:UPI0000051830; GB:AE005672;
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SPI409
C:Superfamily: oxygen-independent coporophyrinogen oxidase

Query Match 12.6%; Score 250.5; DB 2; Length 376;
Best Local Similarity 26.0%; Pred. No. 5.6e-10;
Matches 89; Conservative 57; Mismatches 151; Indels 45; Gaps 10;
Qy 12 SLYNLSFCCK-----DF-----SIFNRILSELKCHLILGHPIIKTYIKHVD 55
Db 7 SAYVHIPFCTQICYCDFSKVFIKNQPVDSYLEHLLSEFRSYDI-----EKLSTLYIGGGT 62
Qy 56 FCLSRQDNLFKFTSLSKYINLELLEFTLEIIPGYVDPEKFLDLDEFCTIRINLVQSF 115
Db 63 PTALSAPOLEVLNGLTKNLDLSVLELTIEANPGDLADKIAVLKNSAVNRVSLGVQTF 122
Qy 116 SLEFRKIVGPIEISYKKNILNINIRKPPDP-LNIDMTVNMPLOKSHLKDRLKELLSYM 174
Db 123 DDKMLKKIGRSHLE-KDIYENIDRLKLAGFNSIDLIYALPGQTMEQVKENAKAIGLD 181
Qy 175 PEHICFSDICEEGFVLDRF-----DNSIDSEKLMFCALECSNGYINVEITNEA 226
Db 182 IPHMSLSLIDENHTVFNMRMRGKLPKPEEAEEMFEYIIAE-LERAGFEHYEISNFS 240
Qy 227 LKGHESRHNKLNELKPHIGLGLYAVSLLEFCNDKNNNR-----ALIRKTSFVKANNHLV 282
Db 241 KPGFESRHNLMYDNEAYYIGAGASGV-----NGVRYKHGPIRHVLSAVEEGNACI 294
Qy 283 TFELLEDELFVYHFYFIQGLGTIQGVSLRALRLRFEYNEKQPF 324
Db 295 TEDHLSQKEQMEEMFGLRKKSGVSM---ARFEKFGQSF 332

RESULT 7
G83817
C:coporophyrinogen iii oxidase hemN [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: G83817
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; UID:20512582; PMID:11058132
A:Accession: G83817
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-385 <STO>
A:Cross-references: UNIPROT:Q9KD75; UNIPARC:UPI00000C3B7C; GB:AF001511; GB:BA000004; NID
A:Experimental source: strain C-125
C:Genetics:
A:Gene: hemN
C:Superfamily: oxygen-independent coporophyrinogen oxidase

Query Match 12.6%; Score 249; DB 2; Length 385;
Best Local Similarity 25.5%; Pred. No. 7.3e-10;
Matches 100; Conservative 60; Mismatches 156; Indels 76; Gaps 17;
Qy 14 YINLSFC-----CKDPSIF-----NRILEELKCH-LILLGHPI--IKTYIKHVD 56
Db 6 YIHIPFCEHCYCYC-DPNKRYLKNQPNVEYLQALETEMAMVVAEQPKYSQTLVGGGTP 64
Qy 57 CLSRQDNLFKFTSLSKYINLELLEFTLEIIPGYVDPEKFLDLDEFCTIRINLVQSF 116

Db	65	TALTADQALQASIKRTPLSLDEFTFEVNPDSIDEEKLDVLSYGVDRLSIGVQAFQ	124
Qy	117	LEFPKIVGIEPISYKKNILINIRKPPF-DLNDMTVNNPQKSKHLDKELLSYMP	175
Db	125	PLLLKEIGRTH-DQKSVEQAVEKSRQAGFANLSLDLMGLGPKQTPPEMAETLKEAFALV	183
Qy	176	EHI-CFSDFICEEGFVLRDPDNGI-----DSKLMFCALECLSNYINYEITNFALK	228
Db	184	EHLSCYSLKVEAKTVFNRRQGRRLTPPEDEDEVMTRQLCYETEKHGFQYELSNPAKK	243
Qy	229	GHESRHNKLANWELKPHLGLGL-----YAVSLLFCDNCKNNVRALIRKTGSFVK	276
Db	244	GYESRNLVYNNDEYVYGFAGAGHYGVGVYNNH-----GPKLYLQAMEEG	291
Qy	277	-----ANNHLVTFELLEDLEFFVHFHQGLGTIOGVSRLRALRPFYN-----EKQFPQF	326
Db	292	RRPVFESHHSRVQMBEQMFL-----GLKRSGVVEVRFVGVSMFSLEYEKQIAQL	345
Qy	327	INYCSTLSKKVFDDNIM-----LLKGRERFK	353
Db	346	VARC-LLERT---DDRVLTDGLLGNVEVFE	373
RESULT 8			
D72288			
hypothetical protein TM1166 - Thermotoga maritima (strain MSB8)			
C:Species: Thermotoga maritima			
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004			
C:Accession: D72288			
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey			
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;			
C.M.			
Nature 399, 323-329, 1999			
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq			
A:Reference number: A72200; MUID:99287316; PMID:10360571			
A:Accession: D72288			
A:Status: preliminary			
A:Molecule type: DNA			
A:Residues: 1-374 <ARN>			
A:Cross-references: UNIPROT:Q9XOP9; UNIPARC:UPI00000D3914; GB:AE001773; GB:AE000512; NID			
A:Experimental source: strain MSB8			
C:Genetics:			
A:Gene: TM1166			
C:Superfamily: oxygen-independent coproporphyrinogen oxidase			
Query Match 11.9%; Score 236; DB 2; Length 374;			
Best Local Similarity 22.6%; Pred. No. 5.5e-09;			
Matches 89; Conservative 68; Mismatches 148; Indels 88; Gaps 12;			
Qy	9	VELSLYNLSFC-----CKDFSI-----FNRILSELKCHLILLGHPIIKTLIYKHV	54
Db	1	MKLAVYHVHPFCKVCVYCDYFVSVSEKDFYFSHLLREIDLYREVLSESEIKTVYFGG	60
Qy	55	DFCLSRQDNLAFFITFSLSKYINLELEBEFTLEIIPGVYDFEKFLLDFECITRINLVQS	114
Db	61	TPSVVPPSFLKWLKLEKLVRSRGFTPDITIEVNPESVETELKIYKQIGNRLSLGVQA	120
Qy	115	FSLEFRKIVG---IPEISYKKNILINIRKPPFDLNDMTVNNPQKSKHLDKELL	171
Db	121	CDDTVLKNAGRLYKEETLKKAKIVLEQFE---NVNFDLILGLPGETDITLKKDFRLE	176
Qy	172	SYMEPHICSPDICE-----EGFVLRDPDNSIDSEKLMFCALECLSNYIN	220
Db	177	BFPQGVH--SLYLLVEBDRTRFLDLIOKGLV--ELPEEDDVERRHDLFVEFLKERGFURY	232
Qy	221	EITNPALKGHESRHNKLANWELKPHLGLGLYAVSLLFCDNCKNNVRALIRKTGVSFKVANNH	280
Db	233	EISNPAKFGKSLHMLFWRNENYLGIGL-----VSAGGH	266
Qy	281	LVTPELL---EDLEPF-----VYHFI-----OGLGTIOGVSRLRALRLR	315
Db	267	IGRFYVNASDLKEYEEKITKGELPYEVVHENTEBEEALETVFMGLRIKGVGLNRVKIL	326

Qy	316	FEYNEKQFFOFINYCSTLSKK-FVFDNIMLLK	347
Db	327	LPLEKLQKYPCLYKVKNGKIFLSEGDGMFSK	359
RESULT 9			
F89939			
hypothetical protein hemN [imported] - Staphylococcus aureus (strain N315)			
C:Species: Staphylococcus aureus			
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004			
C:Accession: F89939			
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguci			
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;			
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.			
Lancet 357, 1225-1240, 2001			
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.			
A:Reference number: A89758; MUID:21311952; PMID:11418146			
A:Accession: F89939			
A:Status: preliminary			
A:Molecule type: DNA			
A:Residues: 1-374 <KUR>			
A:Cross-references: UNIPROT:Q99TR6; UNIPARC:UPI00000D767E; GB:BA000018; PID:gl3701381; P			
A:Experimental source: strain N315			
C:Genetics:			
A:Gene: hemN			
C:Superfamily: oxygen-independent coproporphyrinogen oxidase			
Query Match 11.7%; Score 232; DB 2; Length 374;			
Best Local Similarity 25.9%; Pred. No. 1e-08;			
Matches 101; Conservative 73; Mismatches 142; Indels 74; Gaps 19;			
Qy	12	SLYNLSFCCCK--DPSIFNRILSELKCHLILLGHPIIKTLIYKHVDFCLSRQDNLKFI-	68
Db	5	SAVIHIPFCVRICTYCDFNK-----YFIQNPQVDEYLDALITENSTAKYRNLTMYV	56
Qy	69	-----TSLSKYIN-LELL-----EEFTLEIIPGVYDFEKFLLDFECITRINLN	111
Db	57	GGTPTALS--INQLERLLKAIIRDFTTITGEYFEANPDELTKKVLLEKYGVNRISMG	114
Qy	112	VQSFLSFRKIVG----IPEISYKKNILINIRKPPFDLNDMTVNNPQKSKHLDKRD	167
Db	115	VQTFKPELLSVLGRTHNTEDIYTSVLNAKNAGIK---SISDLMYHLPKQTIEDFEQSL	170
Qy	168	KEILLSYMPHEICSPDICE-----EGFV-LRDPDNSIDSEKLMFCALECLSNY	217
Db	171	DLALDMDIQHISYGLILEPKTQFYNNYRKGLKLPNEDLGADMYQL---LMSKIQSPF	227
Qy	218	INYEITNFALKGHESRHNKLANWELKPHLGLGLYA-----VSLLFCDNCKNNVRALIRKT	271
Db	228	HQYEISNFALDGHSEHKNKYWNNEEYVGFAGAGSYVDGVRYTNINPVNHYIKAINKES	287
Qy	272	GSFVKANNHLVTFELLEDLEFFVHFVHFTIOGLGTIOGVSRLRALRLRPFYN-EKQFPQFINYC	330
Db	288	KAILVSNKPSLT-ERMEE-EMFL-----GLRLNEGVSRRPKKFKDQSIESVFQGTIN--	338
Qy	331	STLSKKVFDDN--IMLLK-----GRERFK	353
Db	339	NLKEKELIVEKNDAIALTKRGKVGIGNEVFE	368
RESULT 10			
H70423			
oxygen-independent coproporphyrinogen III oxidase - Aquifex aeolicus			
C:Species: Aquifex aeolicus			
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004			
C:Accession: H70423			
R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V			
Nature 392, 353-358, 1998			
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.			
A:Reference number: A70300; MUID:98196666; PMID:9537320			
A:Accession: H70423			

A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-370 <AOF>
A;Cross-references: UNIPROT:O67418; UNIPARC:UPI0000056603; GB:AE000739; NID:G2983813; PI
A;Experimental source: strain VF5
C;Genetics:
A;Gene: hemF
C;Superfamily: oxygen-independent coproporphyrinogen oxidase

Query Match 11.2%; Score 222.5; DB 2; Length 370;
Best Local Similarity 25.3%; Pred. No. 4.6e-08;
Matches 86; Conservative 65; Mismatches 156; Indels 33; Gaps 9;

Qy 12 SLYINLSFC-----CKDFSI-FNR-----ILLELKHCHLLGHPIIK--TLYIKHVDFC 57
Db 4 ALYVHVPCSKYKPCYDFLSLVNSPVSVRDYLKREINLYRELDVKLETLYFGGGTPT 63
Qy 58 LSRDNLKFIPTSLSKYINLELLEFTELEIIPGVVDFEKFLLDFECITRINLNVQSFSL 117
Db 64 LLKPEELGEIEEDKVPGLSGVKEIVTECPYRKFEELKGLGNVISGVGFTE 123
Qy 118 EFRKIVGPIRISYKKNLILNIRKPPFDLNDMTVMNPLOKKSHLKRDLKELLSYMPH 177
Db 124 KGLRVLRKHTVKDLSCLSLPAGIENNVNVDLWGPQSLKDLVELEHLKFPVKH 183
Qy 178 ICPSDFICEB-----GFVLRDFNSIDSE--KWFCALECLSGNINYEITNFALKGH 230
Db 184 VSAYLTLTYEDTPMGLAFKRGFKPELEEEVVLHMLSEGLRGLFKRYEISNWAKEGF 243
Qy 231 ESRHKNLNLWELKHGLGLVLA---VSLFCNDKNNVRALIRKTSFVKANNHLVTFEL 286
Db 244 ECRNLVTKMREELFGVGVGANGYENVRNTGN-----ISKYVRKLKEDKKPFVRV 297
Qy 287 -LEDLEFVYHFIOGLGTIOQVSLRALRLRPFNEYKOFFQ 325
Db 298 QLDTELEKERIMGLATTEGIEKYLKFPEY-LRDFFE 336

RESULT 11
B69640
coproporphyrinogen oxidase (EC 1.3.3.3) III, oxygen-independent hemN - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: B69640
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillette, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: B69640
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-366 <KUN>
A;Cross-references: UNIPROT:P54304; UNIPARC:UPI000006075C; GB:Z99117; GB:AL009126; NID:G
A;Experimental source: strain 168
C;Genetics:
A;Gene: hemN
C;Superfamily: oxygen-independent coproporphyrinogen oxidase
C;Keywords: oxidoreductase

Query Match 10.2%; Score 202.5; DB 1; Length 366;
Best Local Similarity 24.7%; Pred. No. 1.1e-06;

Matches 82; Conservative 57; Mismatches 146; Indels 47; Gaps 12;

Qy 12 SLYINLSFC---CK--DFSIF-----NRILEELKHCHLLGHPIIKTLIYIKHVD 55
Db 3 SAVIHIPFCEHICHYCDNFKYFIQSPQVDEYVLALEQEMINTIAKTQPDLTIFIGGGT 62
Qy 56 FCLSRQDNLKFIPTSLSKYINLSEL-LEEFTELEIIPGVVDFEKFLLDFECITRINLNVQS 114
Db 63 PTSLSBEQLKJMDMINRVLKSPSDLSEFAVDANPPDLSAEKULKILKEAGVNLRSFGVQT 122
Qy 115 FSLEFRKIVGPIRISYKKNLILNIRKPPFD-LNDMTVMNPLOKKSHLKRDLKELLSY 173
Db 123 FEDDLLEKIGRVH-KQKOVFTSPERAREIGFENISLDMFLPGQTLKHLEHSINTALS 181
Qy 174 MPBHCFSDFICEBEGFVLDRDFNSIDSEKL-----WFCALCELESNGYINYEI 222
Db 182 DAESHYSVSLIVPQTV---FYNLMQKGRHLHPQOEAEYMEIVMSKMEAHGIHQYEI 237
Qy 223 TNFALKGHESRHNKLNWELKPHLGLGLYAVSLFCNDKNNVRA-----LIRKTSFVK 276
Db 238 SNFAKAGMESKHNLTWNSNEQYFGFAGAHGYI-GGTRTVNVGVKHYIDLIAEKG-FPY 295
Qy 277 ANNHLVTFELLEDLEFVYHFIOGLGTIOQVS 308
Db 296 RDTEVTTTREQIEEMFL-----GLRKTAGVS 322

RESULT 12
D71354
probable oxygen-independent coproporphyrinogen III oxidase - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C;Accession: D71354
R;Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; Mcdon
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: D71354
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-393 <COL>
A;Cross-references: UNIPROT:O83216; UNIPARC:UPI00000D3217; GB:AE001202; GB:AE000520; NID
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0186
C;Superfamily: oxygen-independent coproporphyrinogen oxidase

Query Match 10.2%; Score 201.5; DB 2; Length 393;
Best Local Similarity 23.9%; Pred. No. 1.4e-06;
Matches 89; Conservative 52; Mismatches 132; Indels 99; Gaps 16;

Qy 9 VELSLYINLSFC-----CKDFSI-----FNRILEELKHCHLLGHPIIK 47
Db 5 VGLASLVHPFCAQRCAYCDFYSLVRSTYFRPQCPHFIDRLDQVALQRECFGVQGVG 64
Qy 48 TLV-----IKHVDVFCLSRQDNLKFIPTSLSKYINLELLEFTELEIIPGVVDF 94
Db 65 TVTMGGGTSLAPQDIRH--FCVA-----LRAAQRY-----PQEFTELVNPEVDTE 110
Qy 95 EKPKLDFECITRINLNVQSFSS---LEFRKIVGPIRISYKKNLILNIRKPPFDLNI-- 149
Db 111 EFLCACAGGVNLSLGVQSLRDEVLRRAERASACARTALRVMTANARFFSGGVISA 170
Qy 150 DMVTNMPLOKKSHLKRDLKELLSYMPHEHCFSDFIC-----EEGFLVRDFNSIDSE 202
Db 171 DLIAGLRGQTARMVRVEDIDELLFGLRHVSLYG-LCVPHPTPTQEBRIA----- 218
Qy 203 KWFCALECLSGNINYEITNFA-LKGHSRKNLNLWELKPHLGLGLYAVSLLCNDKN 261
Db 219 ALWAHGSAYLVRAGFNRYELSNFARTAADESAHNRAWYRWAPAGVGVGATRFVNL 278

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:27:53 ; Search time 81.0092 Seconds
(without alignments)
3283.383 Million cell updates/sec

Title: US-10-688-058-12
Perfect score: 1981
Sequence: 1 MRVDLLPLVLSLYINLSFC.....LVKLIHFNDNFVKVLRLP 377

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1981	100.0	377	2	051600 BORBU
2	1716	86.6	377	2	Q66014 BORGA
3	292.5	14.8	365	2	Q8R695 FUSNN
4	280	14.1	375	2	Q8RB71 THEYN
5	277	14.0	382	2	Q89207 CLOTE
6	275.5	13.9	377	2	Q8X158 CLOPE
7	265	13.4	378	2	Q5L2C1 STRT1
8	262	13.2	378	2	Q5M3X3 STRT2
9	258.5	13.0	379	2	Q9CGF7 LACLA
10	256.5	12.9	376	2	Q8P177 STRP8
11	256.5	12.9	398	2	Q878L1 STRP3
12	256.5	12.9	408	2	Q8K7R5 STRP3
13	255.5	12.9	376	2	Q99ZM6 STRP1
14	255.5	12.9	408	2	Q5XCD8 STRP6
15	253	12.8	396	2	Q8YMH7 ANASP
16	251	12.7	374	2	Q97J35 CLOAB
17	250.5	12.6	376	2	Q8DPA6 STRR6
18	250.5	12.6	376	2	Q97Q22 STRPN
19	250	12.6	376	2	Q8E046 STRAS
20	249	12.6	385	2	Q9KD75 BACHD
21	244	12.3	376	2	Q8E583 STRA3
22	239.5	12.1	374	2	Q6GG87 STAAH
23	239.5	12.1	383	2	Q93R30 TETHA
24	237	12.0	375	2	Q5ZT11 LEGPH
25	236	11.9	374	2	Q9X0P9 THEMA
26	236	11.9	376	1	Y532 BUCAP
27	235	11.9	374	1	Y498 BUCBP
28	233.5	11.8	375	2	Q5WU92 LEGPL
29	233.5	11.8	375	2	Q5X2T4 LEGPA
30	232.5	11.7	380	2	Q8D7D3 STRMU
31	232	11.7	374	2	Q6G8Y4 STAAH

32	232	11.7	374	2	Q7A0S0 STAAW
33	232	11.7	374	2	Q7A5C2 STAAH
34	232	11.7	374	2	Q99TR6 STAAH
35	231.5	11.7	385	2	Q5HFH7 STAAH
36	223.5	11.3	375	2	Q83DU5 COXBU
37	223	11.3	388	2	Q5WHG4 BACSK
38	222.5	11.2	370	2	Q67418 AQUAE
39	218.5	11.0	391	2	Q8CXD1 OCEIH
40	215	10.9	349	2	Q6SH51 BACLD
41	213.5	10.8	339	2	Q9FOM0 MYCIO
42	213.5	10.8	349	2	Q9FOM1 MYCIO
43	213.5	10.8	349	2	Q9FOM2 MYCIO
44	213	10.8	375	2	Q8KAY8 CHLTE
45	212	10.7	376	2	Q8A5S0 BACTN

ALIGNMENTS

RESULT 1
OS1600 BORBU PRELIMINARY; PRT; 377 AA.
AC OS1600;
DT 01-JUN-1998 (T-EMBLrel. 06, Created)
DT 01-JUN-1998 (T-EMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Oxygen-independent coproporphyrinogen III oxidase, putative.
GN OrderedLocusNames=BB0656;
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=139;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
Richardson S.L., Peterson J.D., Kervatage A.R., Quackenbush J.,
Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
Gocayne J.D., Weidman J.F., Utterback T.R., Matthey L., McDonald L.A.,
Artach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi."
RL Nature 390:580-586 (1997).
DR EMBL; AE001167; AAC67009.1; -; Genomic_DNA.
DR PIR; G70181; G70181.
DR TIGR; BB0656; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR InterPro; IPR006638; EIP3/MiAB/NiF.
DR InterPro; IPR007197; Radical SAM.
DR Pfam; PF04055; Radical_SAM; I.
DR SMART; SM00729; EIP3; I.
KW Complete proteome.
SQ SEQUENCE 377 AA; 44756 MW; B07859CFC7793561 CRC64;

Query Match	100.0%;	Score	1981;	DB	2;	Length	377;
Best Local Similarity	100.0%;	Pred. No.	9.1e-128;				
Matches	377;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MRVDLLPLVLSLYINLSFCCKDFSFNRIELKCHLILGHPIKTYIKHVDCLSR	60				
Db	1	MRVDLLPLVLSLYINLSFCCKDFSFNRIELKCHLILGHPIKTYIKHVDCLSR	60				
Qy	61	QDNLKFTFTSLSKYINLELEEFTELTBIIPGVDFEFKFLDEFCTRINLNVSFSLEFR	120				
Db	61	QDNLKFTFTSLSKYINLELEEFTELTBIIPGVDFEFKFLDEFCTRINLNVSFSLEFR	120				
Qy	121	KIVGIPISYKKUNILINNIRKPPFDLINDMTVMPLQKSHLKRDLKELLSYMPHFICF	180				
Db	121	KIVGIPISYKKUNILINNIRKPPFDLINDMTVMPLQKSHLKRDLKELLSYMPHFICF	180				

```

Db 121 KIVGPEISYKKNILINNIRKFPFDLNDMTVMNPLQKSHLKDRLKELLSYMPHEICF 180
Qy 181 SDFICEEGFVLRDFDONSIDSEKLPFCALCLESNGYINYEITNFALKGHESRHNKLWE 240
Db 181 SDFICEEGFVLRDFDONSIDSEKLPFCALCLESNGYINYEITNFALKGHESRHNKLWE 240
Qy 241 LKPHLGGLYAVSLFCNDKNNNRALIRKTSFVKANNHLVTFELLEDEFFVYHFIQ 300
Db 241 LKPHLGGLYAVSLFCNDKNNNRALIRKTSFVKANNHLVTFELLEDEFFVYHFIQ 300
Qy 301 LGTIQVSLRALRFRFYNEKQFQFINYCSTLSKFKFVFDNDIMLLKGRERFKLNFYLVK 360
Db 301 LGTIQVSLRALRFRFYNEKQFQFINYCSTLSKFKFVFDNDIMLLKGRERFKLNFYLVK 360
Qy 361 IINHFNDFPKVKLRLP 377
Db 361 IINHFNDFPKVKLRLP 377

RESULT 2
Q660L4_BORGA
ID Q660L4_BORGA PRELIMINARY; PRT; 377 AA.
AC Q660L4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE HemN-related protein.
GN OrderedLocusNames=BG0679;
OS Borrelia garinii.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=29519;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=P81;
RA Gloeckner G., Lehmann R., Romualdi A., Pradella S.,
RA Schulte-Spechtel U., Wilske B., Suehnel J., Platzer M.;
RT "Comparative analysis of the Borrelia garinii genome.";
RL Submitted (JUN-2004) to the ENBL/Genbank/DBJ databases.
DR EMBL; CP000013; AAU07507.1; -; Genomic_DNA.
DR InterPro; IPR006638; Elp3/MiaB/NifB.
DR InterPro; IPR007197; Radical SAM.
DR Pfam; PF04055; Radical SAM; I.
DR SMART; SM00729; Elp3; I.
KW Complete proteome.
SQ SEQUENCE 377 AA; 44968 MW; C8192B7F0CD24051 CRC64;

Query Match 86.6%; Score 1716; DB 2; Length 377;
Best Local Similarity 85.9%; Pred. No. 1.3e-109;
Matches 324; Conservative 27; Mismatches 26; Indels 0; Gaps 0;

Qy 1 MRVDLLPLVELSLYINLSFCCKDFSIENRILEELKCHLILGHPIIKTYIKHVDVCLSR 60
Db 1 MRVDLLPLVELSLYINLSFCCKDFSIENRILEELKCHLILGHPIIKTYIKHVDVCLSR 60
Qy 61 QDNLKPIFTSLSKYINLEELLEFTLEIIPGVYDFEKKPLLDLDFCITRINLVQSFSLEFR 120
Db 61 QDSLKPIFTSLSKYINLALLEEFTLEIIPGVYDFEKKPLLDLDFCITRINLVQSFSLEFR 120
Qy 121 KIVGPEISYKKNILINNIRKFPFDLNDMTVMNPLQKSHLKDRLKELLSYMPHEICF 180
Db 121 KIMGPEISYKKNILINNIRKFPFDLNDMTVMNPLQKSHLKDRLKELLSYMPHEICF 180
Qy 181 SDFICEEGFVLRDFDONSIDSEKLPFCALCLESNGYINYEITNFALKGHESRHNKLWE 240
Db 181 SDFICEEQDLTRDFDSIDSEKLPFCALCLESNGYINYEITNFALKGHESRHNKLWE 240
Qy 241 LKPHLGGLYAVSLFCNDKNNNRALIRKTSFVKANNHLVTFELLEDEFFVYHFIQ 300
Db 241 LKPYLGGLYAVSLFCNDKNNNRALIRKTSFVKANNHLVTFELLEDEFFVYHFIQ 300
Qy 301 LGTIQVSLRALRFRFYNEKQFQFINYCSTLSKFKFVFDNDIMLLKGRERFKLNFYLVK 360

```

```

Db 301 LGTIQVSNLRSRLRLEFYNDKQFFHFYINYSNLSKKFVFDNDIMLLKGRERFKLDFYLLK 360
Qy 361 IINHFNDFPKVKLRLP 377
Db 361 IINHFNDFPKVKLRLP 377

RESULT 3
Q8R695_FUSNN
ID Q8R695_FUSNN PRELIMINARY; PRT; 365 AA.
AC Q8R695;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Oxygen-independent coproporphyrinogen III oxidase (EC 1.-.-.-).
GN OrderedLocusNames=FN0560;
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=ATCC 25586;
RC STRAIN=ATCC 25586;
RX DOI=10.1128/JB.184.7.2005-2018.2002;
RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fongstein M., Kyridides N.C., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586.";
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL; AE009951; AAL94756.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004109; F:coproporphyrinogen oxidase activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0006779; P:porphyrin biosynthesis; IEA.
DR InterPro; IPR006638; Elp3/MiaB/NifB.
DR InterPro; IPR010723; HemN C.
DR InterPro; IPR004559; HemN rel.
DR InterPro; IPR007197; Radical SAM.
DR Pfam; PF06969; HemN C; 1.
DR Pfam; PF04055; Radical SAM; 1.
DR SMART; SM00729; Elp3; I.
DR TIGRFAMs; TIGR00539; hemN_rel; 1.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 365 AA; 43029 MW; 41EEF343A04309D6 CRC64;

Query Match 14.8%; Score 292.5; DB 2; Length 365;
Best Local Similarity 24.9%; Pred. No. 5.2e-12;
Matches 101; Conservative 74; Mismatches 133; Indels 97; Gaps 16;

Qy 14 YINLSPFC--CK--DFS-----IFNRILEELKCHLILGHPIIKTYIKHVDVFC 57
Db 8 YIHPPCERKCNVCDFTSLKGTDSQTEKYINYLKEIE-----IYSKEVDLS 54
Qy 58 LSRQDNLKF-----IFTSLSKYI---NLELLEFTLEIIPGVYDFEKKPLDDEFICIT 106
Db 55 -EKQDTYFGGTPSLLPINSLEKLSKFSYDKNTEITIEVNPKTVDTNKLKYRKLGIN 113
Qy 107 RINLVQSFSLEFRKIVGPEISYKKNILINNIRKFPFDLNDMTVMNPLQKSHLKR 165
Db 114 RLSIGIGITFDNDNLKVLGRHSHSQEAIEV-YNLARECGFDNISLDIMFSLPYQTLMLQN 172
Qy 166 DLKELLSYMPHEICFSDFCIEEGFVLRDF-----DNSIDSEKLPFCALCLESNGY 217
Db 173 DLEKLVSLNPNHISYSLIWEEGTKFFRDLKSGKLKTFDNDLEA-SMVEYIIIEFLSKDY 231
Qy 218 INYEITNFALKGHESRHNKLNLWELKPHLGGLYAVSL-----FCNDKNNV 264
Db 232 IHVEISNFSKDFESHRSNTYWNENKYLGVLSAAGYLNVRKYNFFNLKDYNNLDNRNI 291

```

```
Qy 265 RALIRKGTSGVKANNHLVTFELLEDLEFFVHTFQGLGTQGVSLRALRLRFENNEKQFF 324
Db 292 LPIDEK-----EILTEEDIEQRYLVGFRLNKKIIP-----SEK--- 326
Qy 325 QFINVCSTLSKKFVFDNIMLLKGRFRKLFYLVKLIINHFNDFN 369
Db 327 -YLEKMSLCKEG-----YLLEKENGILSHKGLMLFNDFISNP 364

RESULT 4
Q8RB71_THETN PRELIMINARY; PRT; 375 AA.
AC Q8RB71_2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DB Coproporphyrinogen III oxidase and related FeS oxidoreductases.
GN Name=HemN; OrderedLocNames=TTE0952;
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
NCBI_TaxID=119072;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336; DOI=10.1101/gr.219302;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RA "A complete sequence of the T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
DR ENBL; AEO13060; AAM24208.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0004109; F:coproporphyrinogen oxidase activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:000287; F:magnesium ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006779; P:porphyrin biosynthesis; IEA.
DR InterPro; IPR002453; Beta tubulin.
DR InterPro; IPR006638; Elp3/Miab/NifB.
DR InterPro; IPR010723; HemN_C.
DR InterPro; IPR004559; HemN_rel.
DR InterPro; IPR007197; Radical_SAM.
DR Pfam; PF04055; Radical_SAM; 1.
DR SMART; SM00729; Elp3; 1.
DR TIGRFAMs; TIGR00539; hemN_rel; 1.
DR PROSITE; PS00228; TUBULIN_E_AUTOREG; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 375 AA; 44547 MW; A42BE382DBCF122C CRC64;

Query Match 14.1%; Score 280; DB 2; Length 375;
Best Local Similarity 24.9%; Pred. No. 3.8e-11;
Matches 96; Conservative 72; Mismatches 149; Indels 68; Gaps 13;

Qy 10 ELSLYINLSFCCK-----DFSI-----NRILEELKCH-----LILLGPI 45
Db 3 EIGYIHPFCCKCYCDFNSYAGYHELMEDYLKALLEVKYTSDRSFRVIVIGGTT 62
Qy 46 IKTLVIKHVDCLSRQDNLFIFTSLSKYINLELEEFTEIIPGYVDPEKFLDDEFCI 105
Db 63 PNFLPPSHVERVLS-----EVHKDYNVSRDAETITVNPGLLTEDKLIYKINGI 112
Qy 106 TRINLVNQSFSLFRKIVGIPISYKLNILINN-----IRKPPDLNIDMTVMPLQKSH 162
Db 113 NRISMGLOAFQNRLLVEIG-----RIHTAEDPLENVALVRKYFDNVNDLIYALPTQSFEE 168
Qy 163 LKRDLKLSYMPHCIFSDPFCIEEGFVLKDF--DNSI-----DSEKLMFCALECELN 215
Db 169 WQETLTQVKVLPKPHISTYSILLENITLFGLYKENRLPVVGESEELKMTWHTGFLKSK 228
Qy 216 GYINYEITNFALKGHESRHNKLNWELPHGLGLYAVSLLFCNDKNNNVRALIRKTSFV 275
```

```
Db 229 GYCHYEISNFPALGYQCRHNILYWECKEYLGFGAGAHs--YFEGTRWNVRIEKYIAIL 287
Qy 276 K---ANNHLVTFELLEDLEFFVHTFQGLGTQGVSLRALRLRFENNEKQFF--FNY- 329
Db 288 KRKDAREIINLSPEDKMSFMF---LGLRMKGVCCEEFKRFKFGISMFERYEIFIKYE 344
Qy 330 -----CSTLSKKFVFDNIM 344
Db 345 KMGLIEKDKDCVRLTEKGDVSNVI 369

RESULT 5
Q892Q7_CLOTE PRELIMINARY; PRT; 382 AA.
AC Q892Q7_2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Oxygen-independent coproporphyrinogen III oxidase (EC 1.-.-.-).
GN OrderedLocNames=CTC02034;
OS Clostridium tetani.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
NCBI_TaxID=1513;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Massachusetts / E88;
RX MEDLINE=22457253; PubMed=12552129; DOI=10.1073/pnas.0335851100;
RA Brueggemann H., Baumer S., Fricke W.F., Wierer A., Liesegang H.,
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
RA Gottschalk G.;
RT "The genome sequence of Clostridium tetani, the causative agent of
RT tetanus disease.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
DR ENBL; AEO15943; AAO36537.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0004109; F:coproporphyrinogen oxidase activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:000287; F:magnesium ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR006638; Elp3/Miab/NifB.
DR InterPro; IPR010723; HemN_C.
DR InterPro; IPR004559; HemN_rel.
DR InterPro; IPR007197; Radical_SAM.
DR Pfam; PF06969; HemN_C; 1.
DR Pfam; PF04055; Radical_SAM; 1.
DR SMART; SM00729; Elp3; 1.
DR TIGRFAMs; TIGR00539; hemN_rel; 1.
DR Complete proteome; Oxidoreductase.
SQ SEQUENCE 382 AA; 45516 MW; 52FA3CD797A6FFAA CRC64;

Query Match 14.0%; Score 277; DB 2; Length 382;
Best Local Similarity 26.7%; Pred. No. 6.3e-11;
Matches 108; Conservative 70; Mismatches 146; Indels 80; Gaps 18;

Qy 1 MYVDLLPLVELSYINLSFC---CK--DFSIFNRILBELKCHILLGHPI-----IK 47
Db 1 MKEDYNNFNKNGIYIHPCKCKCYCDFISYKSGSFMKEYTEALSKEIKRECMKKIR 60
Qy 48 TLYI-----KHVDCLSRQDNLFIFTSLSKYINLELEEFTEIIPGYVDPEKF 97
Db 61 SIYIGGTPSPVVPKWEI-----LKETIDNLSKNL-----EFSIANPGVTVEKL 109
Qy 98 KLDFEFCITRINLVQSFS-----LEFRKIVGIPISYK-KLNILINNIRKPPD-LNIDM 151
Db 110 KPFKEGINRISFGLQAWQHLLNKLGRINHIEFLYSYKL-----ARKLGFNDINVDL 163
Qy 152 TVNMPQKKSHLRDLKELLSYMPHI--CFSDFICE-----EEGFVLKDFPNSIDSE 202
Db 164 MFGIPDQKLNWKEWTELEKIVELNPEHISCSYSLIVEGTTPFYKLYKRGK--NLPNEDIER 221
```



```
QY 229 GHSRHNKLNWELKPHGLGLGYAVSLFLFCNDKNNVRALIRKTG---SFVKA----- 277
Db 243 GMESRHNLMYNNDEYGVGAGAGYV-----NGVR--YNRGPIQHYLKAIADGCHAR 294
QY 278 --NNHLVTFELLEDLEFRVYHFIOGLGTIOGVSRLRRLRFEYNEKQFFQFQINVCSTLSK 335
Db 295 LNEEHLTKVEMMEB-EFFL-----GLRKSQVSIK-----RPE--EKFGLSFSDTYGDIVK 342
QY 336 KFVFDDNIMLLKGRE 350
Db 343 KQEDG--LLVKDPE 355

RESULT 8
Q5M3X3 STRT2 PRELIMINARY; PRT; 378 AA.
AC Q5M3X3
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE Coproporphyrinogen III oxidase.
GN Name=hemN; OrderedLocusNames=stui236;
OS Streptococcus thermophilus (strain ATCC BAA-250 / LMG 18311).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=264199;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinquis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulakauskas S., Lapidus A., Goltzman E., Mazur M., Pusch G.D.,
RA Fongstein M., Overbeek R., Kyprides N., Purnelle B., Prozzi D.,
RA Ngui K., Masny D., Hancy F., Burtreau S., Boutry M., Delcour J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
RT bacterium Streptococcus thermophilus.";
RL Nat. Biotechnol. 22:1554-1558(2004).
DR EMBL; CP000023; AAV60870.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0004109; F:coproporphyrinogen oxidase activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006779; P:porphyrin biosynthesis; IEA.
DR InterPro; IPR006638; Elp3/MiaB/NifB.
DR InterPro; IPR010723; HemN_C.
DR InterPro; IPR004559; HemN_rel.
DR InterPro; IPR007197; Radical_SAM.
DR Pfam; PF06969; HemN_C; 1.
DR Pfam; PF04055; Radical_SAM; 1.
DR SMART; SM00729; Elp3; 1.
DR TIGRFAMs; TIGR00539; hemN_rel; 1.
KW Complete proteome.
SQ SEQUENCE 378 AA; 43467 MW; AB25FBB520786230 CRC64;

Query Match 13.2%; Score 262; DB 2; Length 378;
Best Local Similarity 26.7%; Pred. No. 6.6e-10;
Matches 100; Conservative 71; Mismatches 142; Indels 62; Gaps 17;

QY 12 SLYNLSFCCK-----DFS---IFNRILEELKCHLI-----LLGHPPIKTLYKIKHVDCLS 59
Db 7 SAYVHIPCTQICYCDFSKVPKQNPVDDYLRALIREWELSDIKELATLYIGGTPTAI 66
QY 60 RQDNLKFTSLSKYINILEEFTLEIIPGVDFEKFPLDDEFCTITRINLNVSQSLF 119
Db 67 SAEQDYLHSLQKLDLSKLEEFTEIANPGDLTVDKIEVLKGSANVRVSLGVQTFDDKH 126
QY 120 RKIVGIPISYKLNIL--INNIRKFPF--DLNIDMTVMNPKQKSHLKRDLKELLSYMP 176
Db 127 LRQIG--RSHNQAIYESIDSLKSAGFNISIDLIYALPGQTMQVKNVRKALELDIP 183
QY 177 HICFSDFICEEGFVLRDP-----DNSIDSEKLWFCALCLESNGYINVEITNFALK 228
```

```
Db 184 HLSYLSLLEHHTVFMNKRGRKLNLPTEDEAEFMFDYIIQE--LEKNGFPHYEISNFTKP 242
QY 229 GHSRHNKLNWELKPHGLGLGYAVSLFLFCNDKNNVRALIRKTG---SFVKA----- 277
Db 243 GMESRHNLMYNNDEYGVGAGAGYV-----NGVR--YNRGPIQHYLKAIADGCHAR 294
QY 278 --NNHLVTFELLEDLEFRVYHFIOGLGTIOGVSRLRRLRFEYNEKQFFQFQINVCSTLSK 335
Db 295 LNEEHLTKVEMMEB-EFFL-----GLRKSQVSIK-----RPE--EKFGMSFSDTYGDIVE 342
QY 336 KFVFDDNIMLLKGRE 350
Db 343 KQADG--LLVKDPE 355

RESULT 9
Q9CGF7 LACLA PRELIMINARY; PRT; 379 AA.
AC Q9CGF7
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Oxygen-independent coproporphyrinogen III oxidase.
GN Name=hemN; OrderedLocusNames=LL1139;
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=LI1403;
RX MEDLINE=21235186; PubMed=1137471; DOI=10.1101/gr.GR-1697R;
RA Bolotin A., Wincker P., Mauger S., Jallion O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis LI1403.";
RL Genome Res. 11:731-753(2001).
DR EMBL; AE006346; AAK05237.1; -; Genomic_DNA.
DR PIR; C86767; C86767.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0004109; F:coproporphyrinogen oxidase activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006779; P:porphyrin biosynthesis; IEA.
DR InterPro; IPR006638; Elp3/MiaB/NifB.
DR InterPro; IPR010723; HemN_C.
DR InterPro; IPR004559; HemN_rel.
DR InterPro; IPR007197; Radical_SAM.
DR Pfam; PF06969; HemN_C; 1.
DR Pfam; PF04055; Radical_SAM; 1.
DR SMART; SM00729; Elp3; 1.
DR TIGRFAMs; TIGR00539; hemN_rel; 1.
KW Complete proteome.
SQ SEQUENCE 379 AA; 43552 MW; 44F295AD2C29D4FD CRC64;

Query Match 13.0%; Score 258.5; DB 2; Length 379;
Best Local Similarity 22.8%; Pred. No. 1.2e-09;
Matches 91; Conservative 78; Mismatches 133; Indels 97; Gaps 15;

QY 12 SLYNLSFCCK-----DFSIFNRILEELKCHLIILGHPI-----IKT 48
Db 7 SAYVHIPFCSHCICYCDFA-----KVLMTGQPIDAYIESLIEEFSQSFIEKLRT 55
QY 49 LYXKNDVFCLSGRQNLKFIETSLSKYINILEEFTLEIIPGVYDFEKFPLDDEFCTITRI 108
Db 56 IYIGGGTSPVLSAQQLERLLTAIAEQDLVLEEFTEANPGDLSDEVIKVLADSAVNR 115
QY 109 NLNVQSLSLFRKIVGIPISYKLNILINNIRKFPF--DLNIDMTVMNPKQKSHLKRDL 167
Db 116 SLGVQTFNNALKKIGRTHTEVQVYD--SVRELKKAAGFENITIDLIYALPGQTMEMVKS 174
QY 168 KELLSYNPEHICFSDFICEB-----SGFV--LRDFNSIDSEKLWFCALCLESNGY 217
```

```
Db 175 EKFLKALPHVALYSILLEDHTVPMRQRGLRLPSEDKNAD---NVEYIMDILAKNGY 231
Qy 218 INYEITPALKGCHSRNKLWELKPHGLGLYAVSLFCDNKNVNRALIRKTSFVKA 277
Db 232 NHYESVNFGLPGPKSKNITTYWDNEEYVIGAGAGSYL-----AGIR----- 273
Qy 278 NNHLVTFELLEDLEFFVVFHFIQGLGTIGQVSLRALRLRFEYNEKQFFQFNYNCSTLSKKE 337
Db 274 -----YKULGP-----VHYLKAAPTEKRIINEEVLKSKSQIEENFL-----GLRKK- 315
Qy 338 VFDDNIMLLKGRERFKLNF-----YLVKIINH---FND 367
Db 316 ---SGVLVEKFNKFKCSFKELYGQITELINQKLND 351
```

```
RESULT 10
Q8P177 STRP8
ID Q8P177 STRP8 PRELIMINARY; PRT; 376 AA.
AC Q8P177
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE Putative coproporphyrinogen III oxidase.
GN Name=hemN; OrderedLocuNames=spvM18.1022;
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301451;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=2192593; PubMed=11917108; DOI=10.1073/pnas.062526099;
RA Smoot J.C., Barbican K.D., Van Gompel J.U., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
DR EMBL; AE010030; AAL97657.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cyttoplasm; IEA.
DR GO; GO:0004109; F:coproporphyrinogen oxidase activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006779; P:porphyrin biosynthesis; IEA.
DR InterPro; IPR006638; E1p3/MiAB/NiFB.
DR InterPro; IPR010723; HemN_C.
DR InterPro; IPR004559; HemN_rel.
DR Pfam; PF06969; HemN_C.1.
DR Pfam; PF04055; Radical_SAM.
DR SMART; SM00729; E1p3; I.
DR TIGRFAMs; TIGR00539; hemN_rel; 1.
KW Complete proteome.
SQ SEQUENCE 376 AA; 43093 MW; 7236F821BE82BEE5 CRC64;
```

```
Query Match 12.9%; Score 256.5; DB 2; Length 376;
Best Local Similarity 26.1%; Pred. No. 1.6e-09;
Matches 86; Conservative 63; Mismatches 148; Indels 33; Gaps 9;

Qy 12 SLYNLSFCCK-----DFS-----IPNRILEELKCHLILGHPIIKTLVYKHVD 55
Db 7 SAYVHPFCTQICYYCDSKVFQNPVDAYLKALIQBFESYDI---HD-LKTLVIGGT 62

Qy 56 FCLSRQDNLKFIFTSKYINLELEFEFTLEIIPGYVDFEKFKLDFECITRINLVQSF 115
Db 63 PTATATKQLEYLNHLNHLNLDLEEFTEANPGDLTPKIAVLQRSANVRISLGVTQ 122

Qy 116 SLEFRKIVGIPKISYKKNILNINIRKPPF-DLNIDMTVMNPLQCKSHLKDRLKELLSYM 174
```

```
Db 123 NNKQLKQIGRSH-NBQIYSTIANLKTAGPHNTSIDLIYALPGQTLDOVKENVAKALAD 181
Qy 175 PEHICFSDFCBEEGFLRDF-----DNSIDSEKLWFCALCLESNGVINYEITNFA 226
Db 182 IPHLSLYSLLEHHTVFMKORRGKLNLPTEDEAEFEYIIE-MEANGFEHYEISNFT 240
Qy 227 LKGHESRHNKLNWELKPHGLGLYAVSLFCDNKNVNRALIRKTSFGSKANNHLVTFEL 286
Db 241 KPGFESRHNLMYNDVNEVFCGAGAGSYL--NGIRYQNRVPIQHYLKAVEAGNARLNEEV 298
Qy 287 LEDLEFFVYHFIQGLGTIGQVSLRALRLRP 316
Db 299 LRKEEMMEELFLGLRKKTKGVSIQRFQEK 328

RESULT 11
Q878L1 STRP3
ID Q878L1 STRP3 PRELIMINARY; PRT; 398 AA.
AC Q878L1
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE Putative coproporphyrinogen III oxidase.
GN OrderedLocuNames=SpS1180;
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301448;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SSI-1 / Serotype M3;
RX MEDLINE=22683278; PubMed=12799345; DOI=10.1101/gr.1096703;
RA Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,
RA Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,
RA Hayashi H., Hattori M., Hamada S.;
RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
RT large-scale genomic rearrangement in invasive strains and new insights
RT into phage evolution.";
RL Genome Res. 13:1042-1055 (2003).
DR EMBL; BA000034; BAC64275.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cyttoplasm; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004109; F:coproporphyrinogen oxidase activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0006779; P:porphyrin biosynthesis; IEA.
DR InterPro; IPR006638; E1p3/MiAB/NiFB.
DR InterPro; IPR010723; HemN_C.
DR InterPro; IPR004559; HemN_rel.
DR Pfam; PF06969; HemN_C.1.
DR Pfam; PF04055; Radical_SAM.1.
DR SMART; SM00729; E1p3; I.
DR TIGRFAMs; TIGR00539; hemN_rel; 1.
SQ SEQUENCE 398 AA; 45841 MW; 669A0563D1206261 CRC64;
```

```
Query Match 12.9%; Score 256.5; DB 2; Length 398;
Best Local Similarity 25.8%; Pred. No. 1.7e-09;
Matches 85; Conservative 62; Mismatches 150; Indels 33; Gaps 8;

Qy 12 SLYNLSFCCK-----DFS-----IPNRILEELKCHLILGHPIIKTLVYKHVD 55
Db 29 SAYVHPFCTQICYYCDSKVFQNPVDAYLKALIQBFDSY---GIRDLKTLVIGGT 84

Qy 56 FCLSRQDNLKFIFTSKYINLELEFEFTLEIIPGYVDFEKFKLDFECITRINLVQSF 115
Db 85 PTATATKQLEYLNHLNHLNLDLEEFTEANPGDLTPKIAVLQRSANVRISLGVTQ 144

Qy 116 SLEFRKIVGIPKISYKKNILNINIRKPPF-DLNIDMTVMNPLQCKSHLKDRLKELLSYM 174
Db 145 NNKQLKQIGRSH-NBQIYSTIANLKTAGPHNTSIDLIYALPGQTLDOVKENVAKALAD 203

Qy 175 PEHICFSDFCBEEGFLRDF-----DNSIDSEKLWFCALCLESNGVINYEITNFA 226
```

```

Db 204 IPHLSYLSLILEHHTVFMNVRCKNLNPTDLEAEFVEIISE-MEANGFEHYEISNFT 262
Qy 227 LKGHESRHNKLNWELKPHLGLGLYAVSLFCNDKNNNVRALIRKTSFVKANNHLVTFEL 286
Db 263 KPGFESRHNLMYNDVNEVFGCGAGAGYL--NGIRYQNRVPIQHLYKAVEAGNARLNEEV 320
Qy 287 LEDLEFFVYHFIQGLGTIOGVSRLRLR 316
Db 321 LRKEEMEEELFLGLRKKTVGSIORFQEF 350

RESULT 12
ID Q8K7R5_STRP3 PRELIMINARY; PRT; 408 AA.
AC Q8K7R5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative coproporphyrinogen III oxidase.
GN Name=hemN; OrderedLocNames=SpyW3_0673;
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301448;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGA315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206; DOI=10.1073/pnas.152298499;
RA Beres S.B., Sylva G.L., Barban K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Mueser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).
DR EMBL; AE014150; AAM79280.1; -; Genomic_DNA.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0004109; F:coproporphyrinogen oxidase activity; IEA.
DR GO; GO:0005506; F:magnesium ion binding; IEA.
DR GO; GO:0000287; F:iron ion binding; IEA.
DR GO; GO:0006779; F:porphyrin biosynthesis; IEA.
DR GO; GO:0006779; F:porphyrin biosynthesis; IEA.
DR InterPro; IPR006638; ELP3/MiAB/NifB.
DR InterPro; IPR010723; HemN_C.
DR InterPro; IPR004559; HemN_rel.
DR Pfam; PF06969; HemN_C; 1.
DR SMART; SM00729; ELP3; 1.
DR TIGRFAMs; TIGR00539; hemN_rel; 1.
KW Complete proteome.
SQ SEQUENCE 408 AA; 46987 MW; C31113FBA0B1A5AD CRC64;

Query Match 12.9%; Score 256.5; DB 2; Length 408;
Best Local Similarity 25.8%; Pred. No. 1.7e-09;
Matches 85; Conservative 62; Mismatches 150; Indels 33; Gaps 8;

Qy 12 SLYNLSFCCK-----DFS-----IPNRILEELKCHLLGHPKTLTKYIKHVD 55
Db 39 SAYVHPFQICVYCDFSKVFIONQPVDAVKALIQEFDY----GIRDLTKLYIGGTT 94
Qy 56 FCLSRQDNLFKFTSLSKYNILELEFTLIIIPGYVDPEFKLLDDEFCTIRINLVQSF 115
Db 95 PTATKQLEVLNHLERNLDDLEFTTEANFGDLTPEKIAVLQSAVNRISLGVTFF 154
Qy 116 SLFRKIVGIPETISYKKNILNINIRKPPF--DLNIDMTVMNPKQKSHLRDLKELLSYM 174
Db 155 NNKQLKQIGRSH-NEEQIYSTITNLKAAGFNHISIDLIYALPGQTLDOVKENAKALD 213
Qy 175 PEHICFSDPICEEGFVLDRF-----DNSIDSEKLFWCALECSNGYINYEITNFA 226
Db 214 IPHLSYLSLILEHHTVFMNVRCKNLNPTDLEAEFVEIISE-MEANGFEHYEISNFT 272

```

```

Qy 227 LKGHESRHNKLNWELKPHLGLGLYAVSLFCNDKNNNVRALIRKTSFVKANNHLVTFEL 286
Db 273 KPGFESRHNLMYNDVNEVFGCGAGAGYL--NGIRYQNRVPIQHLYKAVEAGNARLNEEV 330
Qy 287 LEDLEFFVYHFIQGLGTIOGVSRLRLR 316
Db 331 LRKEEMEEELFLGLRKKTVGSIORFQEF 360

RESULT 13
ID Q99ZW6_STRPY PRELIMINARY; PRT; 376 AA.
AC Q99ZW6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative coproporphyrinogen III oxidase (EC 1.3.3.3).
GN Name=hemN; OrderedLocNames=SpY1040;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296; DOI=10.1073/pnas.071559398;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663 (2001).
DR EMBL; AE006549; AAK3932.1; -; Genomic_DNA.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0004109; F:coproporphyrinogen oxidase activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006779; P:porphyrin biosynthesis; IEA.
DR InterPro; IPR006638; ELP3/MiAB/NifB.
DR InterPro; IPR010723; HemN_C.
DR InterPro; IPR004559; HemN_rel.
DR InterPro; IPR007197; Radical_SAM.
DR Pfam; PF06969; HemN_C; 1.
DR SMART; SM00729; ELP3; 1.
DR TIGRFAMs; TIGR00539; hemN_rel; 1.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 376 AA; 43111 MW; 737548C68DB358E3 CRC64;

Query Match 12.9%; Score 255.5; DB 2; Length 376;
Best Local Similarity 25.8%; Pred. No. 1.8e-09;
Matches 85; Conservative 62; Mismatches 150; Indels 33; Gaps 8;

Qy 12 SLYNLSFCCK-----DFS-----IPNRILEELKCHLLGHPKTLTKYIKHVD 55
Db 7 SAYVHPFQICVYCDFSKVFIONQPVDAVKALIQEFDY----GIRDLTKLYIGGTT 62
Qy 56 FCLSRQDNLFKFTSLSKYNILELEFTLIIIPGYVDPEFKLLDDEFCTIRINLVQSF 115
Db 63 PTATKQLEVLNHLERNLDDLEFTTEANFGDLTPEKIAVLQSAVNRISLGVTFF 122
Qy 116 SLFRKIVGIPETISYKKNILNINIRKPPF--DLNIDMTVMNPKQKSHLRDLKELLSYM 174
Db 123 NNKQLKQIGRSH-NEEQIYSTITNLKAAGFNHISIDLIYALPGQTLDOVKENAKALD 181
Qy 175 PEHICFSDPICEEGFVLDRF-----DNSIDSEKLFWCALECSNGYINYEITNFA 226
Db 182 IPHLSYLSLILEHHTVFMNVRCKNLNPTDLEAEFVEIISE-MEANGFEHYEISNFT 240
Qy 227 LKGHESRHNKLNWELKPHLGLGLYAVSLFCNDKNNNVRALIRKTSFVKANNHLVTFEL 286

```


Qy 330 C 330
Db 356 C 356

Search completed: January 24, 2006, 19:56:48
Job time : 83.0092 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:30:59 ; Search time 23.2413 Seconds
(without alignments)
1341.095 Million cell updates/sec

Title: US-10-688-058-12
Perfect score: 1981
Sequence: 1 MRVDLLPLVELSLYINLSFCSIFNIRILBELKCHLLIGHPIIKTIYIKHVDCLSR 377

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	863	43.6	175	2	US-09-830-230A-521
2	763	38.5	155	2	US-09-830-230A-522
3	250.5	12.6	376	2	US-09-830-230A-521
4	248.5	12.5	409	2	US-09-107-433-4312
5	196.5	9.9	391	2	US-09-134-001C-3690
6	196.5	9.9	393	2	US-09-107-532A-6801
7	190.5	9.6	397	2	US-09-134-000C-5885
8	181	9.1	382	2	US-09-543-681A-5208
9	173	8.7	394	2	US-09-198-452A-396
10	173	8.7	402	2	US-09-438-185A-382
11	152.5	7.7	387	2	US-09-489-039A-11948
12	150.5	7.6	389	2	US-09-328-352-8088
13	143.5	7.2	506	2	US-09-352-991A-23066
14	142.5	7.2	401	2	US-09-902-540-14907
15	142	7.2	460	2	US-09-198-452A-959
16	142	7.2	460	2	US-09-438-185A-891
17	140.5	7.1	424	2	US-09-252-991A-20695
18	128	6.5	479	2	US-09-540-236-2194
19	123.5	6.2	5037	2	US-09-424-783-4
20	118.5	6.0	4968	2	US-09-424-783-5
21	118.5	6.0	5032	2	US-09-538-092-979
22	116.5	5.9	487	2	US-09-543-681A-6824
23	116.5	5.9	494	2	US-09-489-039A-12255
24	113.5	5.7	491	2	US-09-248-796A-19312
25	113.5	5.7	701	2	US-09-710-279-710
26	113.5	5.7	703	2	US-09-134-001C-3015
27	113.5	5.7	1381	2	US-09-662-254B-20

28	112.5	5.7	309	2	US-09-830-230A-121	Sequence 121, App
29	110	5.6	471	2	US-09-328-352-5145	Sequence 5145, App
30	109.5	5.5	871	2	US-09-792-024-81	Sequence 81, Appl
31	108.5	5.5	913	2	US-09-248-796A-15734	Sequence 15734, A
32	106	5.4	369	2	US-09-543-681A-7314	Sequence 7314, Ap
33	106	5.4	904	2	US-09-198-484-2	Sequence 2, Appli
34	105.5	5.3	287	2	US-09-830-230A-122	Sequence 122, Appl
35	105.5	5.3	471	1	US-08-477-451-20	Sequence 20, Appl
36	105.5	5.3	618	2	US-09-299-378-4	Sequence 4, Appli
37	104	5.2	899	2	US-09-248-796A-15467	Sequence 15467, A
38	104	5.2	2026	2	US-09-487-558B-86	Sequence 86, Appl
39	104	5.2	4866	2	US-09-424-783-2	Sequence 2, Appli
40	104	5.2	4872	2	US-09-424-783-3	Sequence 3, Appli
41	103.5	5.2	387	2	US-09-248-796A-22214	Sequence 22214, A
42	102.5	5.2	425	2	US-09-134-001C-5654	Sequence 5654, Ap
43	102.5	5.2	1226	1	US-08-540-804-12	Sequence 12, Appl
44	102.5	5.2	1226	1	US-08-218-265-12	Sequence 12, Appl
45	102.5	5.2	1226	2	US-08-521-872-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-09-830-230A-521
; Sequence 521, Application US/09830230A
; Patent No. 6902893
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481US
; CURRENT APPLICATION NUMBER: US/09/830,230A
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: PCT/US98/12718
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/057,483
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 521
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-830-230A-521

Query Match 43.6%; Score 863; DB 2: Length 175;
Best Local Similarity 98.2%; Pred. No. 8.2e-80;
Matches 167; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 MRVDLLPLVELSLYINLSFCSIFNIRILBELKCHLLIGHPIIKTIYIKHVDCLSR 60
Db 1 MRVDLLPLVELSLYINLSFCSIFNIRILBELKCHLLIGHPIIKTIYIKHVDCLSR 60
QY 61 QNCLKPIFTSLSKYINLELLEEFTELEIPGYVDYDFEKLDDFCITRINMLNVQSFSLR 120
Db 61 QNCLKPIFTSLSKYINLELLEEFTELEIPGYVDYDFEKLDDFCITRINMLNVQSFSLR 120
QY 121 KIVGPEISYKKILNINIRKPPFDNLNMTVMNPLQKKSHLKRDLKEL 170
Db 121 KIVGPEISYKKILNINIRKPPFDNLNMTVMNPLQKKSHLKRDLQRI 170

RESULT 2
US-09-830-230A-522
; Sequence 522, Application US/09830230A
; Patent No. 6902893
; GENERAL INFORMATION:

APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Lyme Disease Vaccines
FILE REFERENCE: PB481US
CURRENT APPLICATION NUMBER: US/09/830,230A
CURRENT FILING DATE: 2001-09-27
BEST LOCAL SIMILARITY: 98.0%; Pred. No. 1.1e-69;
PCT/US98/12718
PRIOR APPLICATION NUMBER: 60/057,483
PRIOR FILING DATE: 1997-09-03
PRIOR APPLICATION NUMBER: 60/053,344
PRIOR FILING DATE: 1997-07-22
PRIOR APPLICATION NUMBER: 60/053,377
PRIOR FILING DATE: 1997-07-22
PRIOR APPLICATION NUMBER: 60/050,359
PRIOR FILING DATE: 1997-06-20
NUMBER OF SEQ ID NOS: 756
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 522
LENGTH: 155
TYPE: PRT
ORGANISM: Homo sapiens
US-09-830-230A-522

Query Match 38.5%; Score 763; DB 2; Length 155;
Best Local Similarity 98.0%; Pred. No. 1.1e-69;
Matches 147; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 21 CKDPSIFNRILEELKCHLLGHPHPIIKTYIKHVDPCLSRQDNLKFTFTSLSKYINLELL 80
Db 1 CKDPSIFNRILEELKCHLLGHPHPIIKTYIKHVDPCLSRQDNLKFTFTSLSKYINLELL 60

Qy 81 EEFTLEIIPGVDFEKEFLDDECITRINLNQVSFSEFKYIGVIGPISYKKNILNINI 140
Db 61 EEFTLEIIPGVDFEKEFLDDECITRINLNQVSFSEFKYIGVIGPISYKKNILNINI 120

Qy 141 RKFPFDLNDMTVMNPLQKKSHLKRDLKEL 170
Db 121 RKFPFDLNDMTVMNPLQKKSHLKRDLQRI 150

RESULT 3
US-09-583-110-3497
Sequence 3497, Application US/09583110
Patent No. 669703
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al.
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
FILE REFERENCE: PATH00-07A
CURRENT APPLICATION NUMBER: US/09/583,110
CURRENT FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 3497
LENGTH: 376
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-583-110-3497

Query Match 12.6%; Score 250.5; DB 2; Length 376;
Best Local Similarity 26.0%; Pred. No. 6e-17;
Matches 89; Conservative 57; Mismatches 151; Indels 45; Gaps 10;

Qy 12 SLYINLSPCK-----DF-----STFNRIELKCHLLGHPHPIIKTYIKHVD 55
Db 7 SAYVHIPFCTQICYCDPSKVFIRKNQPVDSYLEHLEEFRSYDI-----EKLSTLYIGGGT 62

Qy 56 FCLSRQDNLKFTFTSLSKYINLELLEPFTLEIIPGVDFEKEFLDDECITRINLNQVSF 115

Db 63 PTALSAPQLEVLNGLTKNLDLSVLEELTEANPGDLDADKIAVLKNSAVNRVSLGVQTF 122
Qy 116 SLEFRKIVGIPISYKKNILNINNRKFPFD-LNIDMTVMNPLQKKSHLKRDLKELLSYM 174
Db 123 DDKMLKKIGRSHLE-KDIYENIDRLKAGFDNISIDLIYALPGQTWEQVKENAKAIGLD 181
Qy 175 PHEICPSDFICEEGEVLRFD-----DNSIDSEKLMFCALECLESNGYINVEITNFA 226
Db 182 IPHMSLYSILLENHTVFMRMGRGKLPKPEELEAEFVYIAE-LERAGFEHYEISNFS 240
Qy 227 LKGHESRHNKLANWELPHGLGLYAVSLFCNDKNNVR-----ALIRKTGTSFYKANNHLV 282
Db 241 KPGFESRHNLMYWDNAEYVIGAGASGYV-----NGVRYKNGHPIRHYLSAVEEGNACI 294
Qy 283 TFELLEDLEFFVYHFQGLGTTCQVSLRALRLRFEYNEKQFF 324
Db 295 TEDHLSQKEQMEEMFLGRKKSQVSM-----ARFEKFGQSF 332

RESULT 4
US-09-107-433-4312
Sequence 4312, Application US/09107433
Patent No. 6800744
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSER: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/ 085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4312:
SEQUENCE CHARACTERISTICS:
LENGTH: 409 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...409
SEQUENCE DESCRIPTION: SEQ ID NO: 4312:
US-09-107-433-4312

Query Match 12.5%; Score 248.5; DB 2; Length 409;


```
; LOCATION: 1...460
; OTHER INFORMATION: Xaa=unknown or other
US-09-198-452A-959

Query Match      7.2%; Score 142; DB 2; Length 460;
Best Local Similarity 18.3%; Pred. No. 9e-06;
Matches 67; Conservative 68; Mismatches 162; Indels 70; Gaps 13;

Qy 11 LSLYINLSFC-----CKDPSIFNR-----ILEELKCHLILLG-HPIIKTLYIK 52
    ||| : : || : : || : : || : : || : : || : : || : : || : : ||
Db 52 LSLYHIPFCQSMCLYCGCSVVLNRREDIVEAYINTLIQEMKLVETIGFRPQVSRHFG 111
    ||| : : || : : || : : || : : || : : || : : || : : || : : ||
Qy 53 HVDFCLSRQDNLKFIPTSLSKYINLELLEETLEIIPGV--DEEKKLLDEFCTIRNL 110
    ||| : : || : : || : : || : : || : : || : : || : : || : : ||
Db 112 GGTPSRLSRELTFLFDHIHKLFDSHAEETAEVDPRLRNDMEKADFFQNVGFNRVSL 171
    ||| : : || : : || : : || : : || : : || : : || : : || : : ||
Qy 111 NVQSFSLFRKIVGIPISYKKNILINNIRKPPF-DLNIDMTVMPLQKKSHLKRDLKE 169
    ||| : : || : : || : : || : : || : : || : : || : : || : : ||
Db 172 GVQDTQADVQEAVERRQ-SHEESLKAYEKELAFQSNINIDLIYGLPKQTKESFSKTIQD 230
    ||| : : || : : || : : || : : || : : || : : || : : || : : ||
Qy 170 LLSTMPHICFSDPIC-----EEGFLVLRDFDINSIDSEKLMFCALECLSENGY----- 217
    ||| : : || : : || : : || : : || : : || : : || : : || : : ||
Db 231 ILAMPDRLALFSPASVPWIKPHQKAMKASDMPSEKFAIYSQSRHLLTKAGYQALGMD 290
    ||| : : || : : || : : || : : || : : || : : || : : || : : ||
Qy 218 ---INYEITWPAKCHESRNKLNWELKPH---LGLGLYAVSL-----FCNDKNN 262
    ||| : : || : : || : : || : : || : : || : : || : : || : : ||
Db 291 HFSLPDPLTLAFKNKTLIRNFQYSLPPEEDLGLGWTSTSFIRGIYLNQAKTLEEYHN 350
    ||| : : || : : || : : || : : || : : || : : || : : || : : ||
Qy 263 NVRALIRKTSFVKANNHLVTFELLEDELEFPVYHFIOGLGTIOGVSRLRLRFEVNEKQ 322
    ||| : : || : : || : : || : : || : : || : : || : : || : : ||
Db 351 TV-----LRGTFTVKSILT-----ED-----DRIRKWAHKLMCTFTINKEE 389
    ||| : : || : : || : : || : : || : : || : : || : : || : : ||
Qy 323 FFQFINY 329
    ||| : : || : : || : : || : : || : : || : : || : : || : : ||
Db 390 FFNLFGY 396
```

Search completed: January 24, 2006, 20:02:42
Job time : 24.2413 secs

• This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:57:15 ; Search time 66.2599 Seconds
(without alignments)
2377.329 Million cell updates/sec

Title: US-10-688-058-12

Perfect score: 1981

Sequence: 1 MRVDLLPLVLSLYINLSFC.....LVKIHNFNDNFFKVKLRLP 377

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA_Main*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*

2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp.*

3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pcp.*

4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp.*

5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp.*

6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1981	100.0	377	5	US-10-688-058-12
2	1090	55.0	204	5	US-10-688-058-10
3	863	43.6	175	5	US-10-994-726-521
4	763	38.5	155	5	US-10-994-726-522
5	250.5	12.6	376	5	US-10-472-928-2860
6	248.5	12.5	409	5	US-10-617-320-4312
7	231.5	11.7	385	5	US-10-470-048B-106
8	206.5	10.4	381	5	US-10-650-274-98
9	173	8.7	394	4	US-10-289-762-396
10	168	8.5	352	4	US-10-335-977-6718
11	168	8.5	364	4	US-10-335-977-6719
12	165	8.3	345	4	US-10-335-977-6717
13	150.5	7.6	493	4	US-10-425-115-261185
14	142	7.2	460	4	US-10-289-762-959
15	135.5	6.8	819	4	US-10-282-122A-47000
16	129.5	6.5	1034	5	US-10-732-923-23287
17	128	6.5	1219	5	US-10-732-923-23209
18	124.5	6.3	548	4	US-10-724-972A-7488
19	124.5	6.3	605	4	US-10-282-122A-54636
20	123	6.2	665	4	US-10-437-963-172583
21	123	6.2	701	4	US-10-282-122A-47140
22	121	6.1	605	4	US-10-238-075-691
23	120.5	6.1	2055	4	US-10-276-774-1795
24	118.5	6.0	4767	4	US-10-276-774-1902
25	118.5	6.0	4967	4	US-10-668-767-60
26	118.5	6.0	4967	4	US-10-668-767-62
27	118.5	6.0	4968	4	US-10-668-767-61

28	118.5	6.0	5032	4	US-10-408-765A-26	Sequence 26, Appl
29	118.5	6.0	5038	5	US-10-723-860-1310	Sequence 1310, Ap
30	118.5	6.0	5038	5	US-10-756-149-5039	Sequence 5039, Ap
31	118.5	6.0	5081	4	US-10-276-774-1850	Sequence 1850, Ap
32	118.5	6.0	5081	5	US-10-450-763-35328	Sequence 35328, A
33	113.5	5.7	703	4	US-10-724-972A-33982	Sequence 3982, Ap
34	113.5	5.7	1381	5	US-10-706-635-20	Sequence 20, Appl
35	112.5	5.7	309	5	US-10-994-726-121	Sequence 121, Appl
36	112	5.7	1478	5	US-10-732-923-3353	Sequence 3353, Ap
37	111.5	5.6	742	4	US-10-437-963-109290	Sequence 109290,
38	111.5	5.6	950	4	US-10-282-122A-47285	Sequence 47285, A
39	111.5	5.6	2326	5	US-10-732-923-22500	Sequence 22500, A
40	110	5.6	1935	5	US-10-732-923-13992	Sequence 13992, A
41	109.5	5.5	871	4	US-10-032-585-7050	Sequence 7050, Ap
42	109.5	5.5	871	5	US-10-882-104-81	Sequence 81, Appl
43	109	5.5	485	4	US-10-282-122A-52735	Sequence 52735, A
44	109	5.5	933	4	US-10-282-122A-47180	Sequence 47180, A
45	108.5	5.5	1032	4	US-10-607-631-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-10-688-058-12
; Sequence 12, Application US/10688058
; Publication No. US20050058661A1
; GENERAL INFORMATION:
; APPLICANT: SYKES, KATHERINE F.
; APPLICANT: HALE, KATHERINE S.
; APPLICANT: JOHNSTON, STEPHEN A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF THE GENUS
; TITLE OF INVENTION: BORRELIA
; FILE REFERENCE: MCRO:00305
; CURRENT APPLICATION NUMBER: US/10/688,058
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-688-058-12

Query Match		100.0%;	Score 1981;	DB 5;	Length 377;
Best Local Similarity		100.0%;	Pred. No. 8.5e-169;		
Matches 377;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MRVDLLPLVLSLYINLSFCCKDFSTFNRILEELKCHLILGHPILIKTLVKKHVDVDFCLSR	60		
Db	1	MRVDLLPLVLSLYINLSFCCKDFSTFNRILEELKCHLILGHPILIKTLVKKHVDVDFCLSR	60		
Qy	61	QDNLKFTISLSKYINLELEEEFTLEIIPGVDFPEFKLLDFECITRINLVQSFSLEFR	120		
Db	61	QDNLKFTISLSKYINLELEEEFTLEIIPGVDFPEFKLLDFECITRINLVQSFSLEFR	120		
Qy	121	KIVGPIEISYKGLNLIINNIRKFPFDLIDMTVMNPLQKSHLKRDLKELLSYMPHEICF	180		
Db	121	KIVGPIEISYKGLNLIINNIRKFPFDLIDMTVMNPLQKSHLKRDLKELLSYMPHEICF	180		
Qy	181	SDFCIEBEGFVLDFDNDSEKLFWFCALCSLNGYINTEITNFALKGHSHRNKLWE	240		
Db	181	SDFCIEBEGFVLDFDNDSEKLFWFCALCSLNGYINTEITNFALKGHSHRNKLWE	240		
Qy	241	LKPHLGLYAVSLFCNDKNNVRALIRKTSFVKANNHLVTFELLEDLEFFVYHFIQG	300		
Db	241	LKPHLGLYAVSLFCNDKNNVRALIRKTSFVKANNHLVTFELLEDLEFFVYHFIQG	300		
Qy	301	LGTTQGVSLRALRFRFYNKQFFQFINYCSTLSKFKVFDDNIMLLKGRBFKLNFLVVK	360		
Db	301	LGTTQGVSLRALRFRFYNKQFFQFINYCSTLSKFKVFDDNIMLLKGRBFKLNFLVVK	360		

```
Qy 361 IINHFNDFKVKLRP 377
Db 361 IINHFNDFKVKLRP 377

RESULT 2
US-10-688-058-10
; Sequence 10, Application US/10688058
; Publication No. US20050058661A1
; GENERAL INFORMATION:
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: HALE, KATHERINE S.
; APPLICANT: JOHNSTON, STEPHEN A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF THE GENUS
; TITLE OF INVENTION: BORRELIA
; FILE REFERENCE: MCRO:003US
; CURRENT APPLICATION NUMBER: US/10/688,058
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-688-058-10

Query Match 55.0%; Score 1090; DB 5; Length 204;
Best Local Similarity 100.0%; Pred. No. 2.5e-89;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 141 RKPPFDLNIDMTVMPLQKSHLKRDLKELLSYMPHEICFSDFCIEBEGFVLRDFDINSID 200
Db 1 RKPPFDLNIDMTVMPLQKSHLKRDLKELLSYMPHEICFSDFCIEBEGFVLRDFDINSID 60

Qy 201 SEKLWFCALCELENGYINYEITTFALKGHESRNKNLWELKPHLGLGLYAVSLFCNDK 260
Db 61 SEKLWFCALCELENGYINYEITTFALKGHESRNKNLWELKPHLGLGLYAVSLFCNDK 120

Qy 261 NNNVALIRKTGSPVKANNHLVTPELLEDLFFVYHFTQGLGTIQGVSLRALRFRFYNE 320
Db 121 NNNVALIRKTGSPVKANNHLVTPELLEDLFFVYHFTQGLGTIQGVSLRALRFRFYNE 180

Qy 321 KQFFQFINCYSTLSKKFVDDNIM 344
Db 181 KQFFQFINCYSTLSKKFVDDNIM 204

RESULT 3
US-10-994-726-521
; Sequence 521, Application US/10994726
; Publication No. US20050147999A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481D1
; CURRENT APPLICATION NUMBER: US/10/994,726
; CURRENT FILING DATE: 2004-11-23
; PRIOR APPLICATION NUMBER: 09/830,230
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: PCT/US98/12718
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/057,483
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 521
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-994-726-521

Query Match 38.5%; Score 763; DB 5; Length 155;
Best Local Similarity 98.0%; Pred. No. 3.4e-60;
Matches 147; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 21 CKDPSIFNRIELKCHLIIILGHPIIKTYIKHVDVFCLSRODNLKFIPTSLSKYINLELL 80
Db 1 CKDPSIFNRIELKCHLIIILGHPIIKTYIKHVDVFCLSRODNLKFIPTSLSKYINLELL 60

Qy 81 BEFTLEIIPGVDPFEKFLDDEFICITRINLNVSQFSLEFRKIVGIPETSYKKLNLINNI 140
Db 61 BEFTLEIIPGVDPFEKFLDDEFICITRINLNVSQFSLEFRKIVGIPETSYKKLNLINNI 120

Qy 141 RKPPFDLNIDMTVMPLQKSHLKRDLKEL 170
Db 121 RKPPFDLNIDMTVMPLQKSHLKRDLQRI 150

RESULT 5
US-10-472-928-2860
; Sequence 2860, Application US/10472928
; Publication No. US20050020813A1
```



```
; SEQ ID NO 106
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-470-048B-106

Query Match      11.7%; Score 231.5; DB 5; Length 385;
Best Local Similarity 24.3%; Pred. No. 4.7e-12;
Matches 94; Conservative

Qy 12 SLVNLSPCK-----DFSPF-----NRILEKCHLLGHPIKTYIKHVPCL 59
Db 16 SAYTHIPFCVRICTYCDNFYIQNPVDSYLDALITETAKYRILKTVGGGPTAL 75
Qy 60 RQMLKFIPTSLSKYINLELEBEPTLEIIPGYVDFEKFLLDFCITRINLVQSPSLEF 119
Db 76 SINQLERLLKAIRD--TFTTITGETTFEANDUKEKVQLEKYGKVRISNGVQTFAPEL 133
Qy 120 RKIVG----IPEISYKKNILNINIRKFPFDMIDTMVNMPLQKSHLKRDLKELLSYMP 175
Db 134 LSVLGRTHNTEDIYTVLNAKNAGIK----SISLDLMYHLPKQTIEDFEQSLDALDMDI 189
Qy 176 EHICFSFICE-----EEGFV-LRDFDINSIDSEKLMFCALECSNGYINYEITNF 225
Db 190 QHISSYGLILEPKTQFYNNMYRKGLLKPNDGLGADMYQL---LMSKIEQSPFHQYEISNF 246
Qy 226 ALKGHESPHKNLWELKPHGLGLGYA-----VSLLEFCNDKNNVRALIRKTSFVKANN 279
Db 247 ALDGHSEHNKVWFNEEYVGFGAGASYVDGVRYYTNPNVNHVIKAINKESKAILVSNK 306
Qy 280 HLVTFFELLEDFPVYHFIOQLGTIQGVSLRALRLRPEYN-EKQFFQFINYCSTLSKKFV 338
Db 307 PSLT-ERMEE-EMFL-----GLRLNEGVSRRPKKFDQSIESVFGQTIN--NLKEKELI 357
Qy 339 FDDN-IMLKGR-----ERPKLN 355
Db 358 VEKNDVIALNTRGKVGIGNEVFEAPLIN 384

RESULT 8
US-10-650-274-98
; Sequence 98, Application US/10650274
; Publication No. US20050202437A1
; GENERAL INFORMATION:
; APPLICANT: GLENN, MATTHEW
; APPLICANT: HAVOKKALA, ILKKA J
; APPLICANT: LUBBERS, MARK WILLIAM
; APPLICANT: DEKKER, JAMES
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES,
; TITLE OF INVENTION: MATERIALS INCORPORATING THEM, AND METHODS FOR USING
; TITLE OF INVENTION: THEM,
; FILE REFERENCE: 11000.1073
; CURRENT APPLICATION NUMBER: US/10/650,274
; CURRENT FILING DATE: 2003-08-28
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: FASTSEQ FOR WINDOWS VERSION 4.0
; SEQ ID NO 98
; LENGTH: 381
; TYPE: PRT
; ORGANISM: LACTOBACILLUS RHAMNOSUS
US-10-650-274-98

Query Match      10.4%; Score 206.5; DB 5; Length 381;
Best Local Similarity 24.7%; Pred. No. 8e-10;
Matches 89; Conservative 63; Mismatches 121; Indels 87; Gaps 17;

Qy 14 YINLSFC---CKDFSIFNRI-----LEELKCHLILGHPI--IIKTLXI--- 51
Db 5 YIHPFCEHICY-YCFDNKVIEGQPVDDYVAMLLKEMR--MVMAEPKEIEIVYVGGG 61
Qy 52 -----KHVDFCLSRQDNLIKFIPTSLSKYINLELEBEPTLEIIPG-YVDFEKFLLDE 102
Db 62 TPTTLTPQLAVALCQGIIRDILHFDPH-----GEFTFEANFNDLLTTDKLQVLYD 109
```

```
Qy 103 FCIITRINLVQSFSLFRKIVGIPEISYKKNILNINIRKFPFD-LNIDMTVMPLQKKS 161
Db 110 FGVRRLSIGVQSFNDVDVLRIGRIHRA-KDVYTAIGNARKVGFNDLSIDLFLRDPQSRD 168
Qy 162 HLKRDLEKLLSYMPEHICFSDFICEBEGFVLRDFDINSIDSEKL-----WFCAL 210
Db 169 DFLNSLQKALALDPLHYSYSLILERTKI----FYNLMRQKRLRPTQDVEADMYQDAID 224
Qy 211 CLESNGVINVEITNFALKGHESPHKNLWELKPHGLGLGYAVSLLFCNDKNNVRALIRK 270
Db 225 LMEAGRHQYEISNFNFAKTGYQCHRNLLYWNQDKYFGGAGAFYL-GRDRHYHYGPIKQY 283
Qy 271 TGSF-----VKANNHLVTFELLEDFEYFVYHFIOQLGTIQGVSLRALRLRPEYNEKQFF 324
Db 284 LAPLHADHLVLAHLVPVSEQIEE-EMFL-----GLRTWAGV-----NEDRFY 326

RESULT 9
US-10-289-762-396
; Sequence 396, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 396
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-289-762-396

Query Match      8.7%; Score 173; DB 4; Length 394;
Best Local Similarity 23.1%; Pred. No. 8.4e-07;
Matches 90; Conservative 68; Mismatches 164; Indels 68; Gaps 16;

Qy 3 VDLLPLV-----ELSYINLSFC-----CKDPSI-FNRILEELKCHLILG-----HPI 45
Db 15 VSLIPMLMGKAPLALYHIHPFCTKKCRYCSFVTIPYKSESVSILYCNNAVIOEGLRKLAPI 74
Qy 46 IKTYLKHVDFP-----CLSRQDNLIKFIPTSLSKYINLELEBEPTLEIIPGVDFEKFLL 100
Db 75 QETHFIETVFFGGTSPSLVPLDLKRLKELAPH-----AREITLEANPENLTVSYLRQL 129
Qy 101 DEFCTRINLVQSFSLFRKIVGIPEISYKKNILNINIRKFPFD-LNIDMTVMPLQK 159
Db 130 QETPINRISGVQTFDDSIQLLGRTHSSAAITAL-QECONHGFSNLSIDLIIYGLPTQS 188
Qy 160 KSHLKRDLKELLSYMPEHICFSDP-----ICEEGFVLRDFDINSIDS 201
Db 189 LEIFLSDLHQALPTIITHISLYNLTDPTHSFYKHKRILVPTIAQEE--ILAEM--SLLA 244
Qy 202 EKLWFCALBCECSNGYINYEITNFALKGHESPHKNLWELKPHGLGLGYAVSLLFCNDKN 261
Db 245 ENL-----LLSQGFQRYELASYAKPDYPAKNHLYYTWDTPRPFGLGVSASQYLH-GERS 296
Qy 262 NN-----VRALIRKTSFVKANNHLVTFELLEDFEYFVYHFIOQLGTIQGVSLRALRL 314
Db 297 KNTSHISHYLURA-VRKNLPTQETSEILPKKERIKKALALRLLEGADLAEFPSTLISML 355
Qy 315 RFEYNEKQFFQFINYCSTLSKK-FVFDDNI 343
Db 356 TDVQLQNLFVHGQCCLANRQGRLFHDTI 385
```

```
RESULT 10
US-10-335-977-6718
; Sequence 6718, Application US/10335977
```

```

Db      283 LKQVETLTKQDKCKLEKFLGLRCCLGVSLSDLDENKVKFLIENKAFIK--NNRLI 337

RESULT 11
US-10-335-977-6719
; Sequence 6719, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
;
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
;
; INFORMATION FOR SEQ ID NO: 6719:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...364
; SEQUENCE DESCRIPTION: SEQ ID NO: 6719:
;
US-10-335-977-6719

Query Match      8.5%; Score 168; DB 4; Length 364;
Best Local Similarity 23.08; Pred. No. 2.le-06;
Matches 82; Conservative 55; Mismatches 113; Indels 106; Gaps 15

Qy      9 VELSLYINLSFC-----CKDFSIFNR-----ILBELKCHL-----ILIGH 43
      : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      18 INMLYIHPFCENKCGYCAFNSENKGLKEEYIQALCLDLKHALSQTDEPIESIFGG 77
      : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

Qy      44 PIKTLYIKHVPCLSRQDNLKIFTSLSKYINLELEBFTLEIIPGVVDEPEKFLDDEF 103
      : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      78 GTENTLSVK-----AFERIFESIHQHASLMDCEITTEANPELLISKAWCQGLKDL 127
      : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

Qy      104 CITRINLVNQSFLSEFAKIVGIPISYKKNILINNIRKPPFD--LNIDMTVMPLQKSH 162
      : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      128 GINRLSIGVQSFPED--KLFLERQHSKIAPVIEIIFKSGIENISDLYINTPLDNETS 185
      : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

Qy      163 LKRDLLKELLSYMP--EHICPSDFICEERGFVLR-----DFDMSIDSEKLWFCALCLE 213
      : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      186 LKELK--LAKELPINHLISAYALSIEKTNLEKNKAKPSSVDFDNVV-----REVLE 235
      : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:21:03 ; Search time 38.8844 Seconds
(without alignments)
1966.133 Million cell updates/sec

Title: US-10-688-058-14
Perfect score: 875
Sequence: 1 QIQTHILASILESIIFILR.....IADTTVHRIIEDGVYLIRTE 174

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	875	100.0	174	8 ADU98755	Adu98755 Borrelia
2	872	99.7	1277	8 ADU98757	Adu98757 Borrelia
3	109	12.5	1629	4 AAU48711	Aau48711 Propionib
4	109	12.5	1629	6 ABM45230	Abm45230 Propionib
5	104.5	11.9	337	7 ADC01324	Adc01324 Enteroha
6	102	11.7	901	9 ADW16234	Adw16234 Mmel endo
7	102	11.7	919	8 ADI61811	Adi61811 Mmel rest
8	99	11.3	1032	6 ABU25689	Abu25689 Protein e
9	97	11.1	374	6 AAE30333	Aae30333 Human pro
10	97	11.1	389	8 ADL23826	Adl23826 Pyrococcu
11	97	11.1	1005	8 ADQ18134	Adq18134 Human sof
12	97	11.1	1113	7 ADN39946	Adn39946 Cancer/an
13	97	11.1	1337	6 ABU56496	Abu56496 Lung canc
14	97	11.1	1337	6 ABU56690	Abu56690 Lung canc
15	97	11.1	1337	7 ADN39060	Adn39060 Cancer/an
16	97	11.1	1359	7 ADN39948	Adn39948 Cancer/an
17	97	11.1	1448	6 ABU56689	Abu56689 Lung canc
18	97	11.1	1448	6 ABU56495	Abu56495 Lung canc
19	97	11.1	1448	7 ADN39058	Adn39058 Cancer/an
20	97	11.1	1448	7 ADN39941	Adn39941 Cancer/an
21	97	11.1	1456	6 ABU56688	Abu56688 Lung canc
22	97	11.1	1456	6 ABU56494	Abu56494 Lung canc
23	97	11.1	1456	7 ADN39943	Adn39943 Cancer/an
24	97	11.1	1456	7 ADN39056	Adn39056 Cancer/an

25	97	11.1	1493	7 ADN39947	Adn39947 Cancer/an
26	97	11.1	1494	6 ABUS6691	Abu56691 Lung canc
27	97	11.1	1494	6 ABUS6497	Abu56497 Lung canc
28	97	11.1	1494	7 ADN39062	Adn39062 Cancer/an
29	97	11.1	1494	7 ADN39945	Adn39945 Cancer/an
30	97	11.1	2308	2 AAR57902	Aar57902 Human RPT
31	97	11.1	2308	4 AAB59586	Aab59586 Human rec
32	97	11.1	2308	6 ABG72433	Abg72433 Recepto
33	97	11.1	2314	6 ABR58521	AbR58521 Human pro
34	97	11.1	2314	6 ABUS6692	Abu56692 Lung canc
35	97	11.1	2314	6 AAE30332	Aae30332 Human PTP
36	97	11.1	2314	7 ADE62111	Ade62111 Human pro
37	97	11.1	2314	7 ADE62042	Ade62042 Human pro
38	97	11.1	2314	7 ADE62038	Ade62038 Cancer/an
39	97	11.1	2314	7 ADN39064	Adn39064 Cancer/an
40	97	11.1	2314	8 ADJ75597	Adj75597 Marker ge
41	97	11.1	2314	8 ADU06410	Adu06410 Novel bro
42	97	11.1	2314	9 ADV70177	Adv70177 Tumor-ass
43	97	11.1	2314	9 AEA00158	Aea00158 Human TAT
44	97	11.1	2314	9 AEA04969	Aea04969 Protein t
45	97	11.1	2314	9 AEA00678	Aea00678 Human TAT

ALIGNMENTS

RESULT 1
ADU98755
ID ADU98755 standard; protein; 174 AA.
AC ADU98755;
XX
XX 24-FEB-2005 (first entry)
XX
DE Borrelia burgdorferi antigenic polypeptide seqid 14.
XX
XX antibacterial; vaccine; immune stimulation; immunity; antigen;
KW DNA library.
XX
OS Borrelia burgdorferi.
XX
XX W02004103269-A2.
XX
XX 02-DEC-2004.
XX
XX 17-OCT-2003; 2003WO-US033056.
XX
XX 18-OCT-2002; 2002US-0419401P.
XX
PA (MACR-) MACROGENICS INC.
PA (TEXA) UNIV TEXAS SYSTEM.
XX
XX Sykes KP, Hale KS, Johnston SA;
XX
XX WPI; 2004-834155/82.
DR N-PSDB; ADU98754.
XX
XX Immunizing a subject against Borrelia burgdorferi infection comprises
XX providing to the subject at least one Borrelia antigen or its fragment.
XX
XX Claim 27; SEQ ID NO 14; 121pp; English.
XX
XX The invention describes a method of immunizing a subject comprising
XX providing to the subject at least one Borrelia antigen or its fragment to
XX induce an immune response. Also described are: an isolated polynucleotide
XX comprising a sequence having at least 17 contiguous nucleotides in common
XX with a sequence not given in the specification; an isolated polypeptide
XX having at least 5 consecutive amino acids of the sequence not given in
XX the specification; a vaccine composition comprising at least one Borrelia
XX antigen or at least one polynucleotide encoding a Borrelia antigen;
XX screening for at least one test polypeptide or test polynucleotide
XX encoding a polypeptide for an ability to produce an immune response;
XX preparing a vaccine; vaccinating a subject; treating a subject infected

CC with a pathogen; passive immunization; vaccination; detecting Herpesvirus
 CC ; and detecting antibodies to a herpesvirus. The method is useful in
 CC immunizing a subject against Borrelia burgdorferi infection. This is the
 CC amino acids sequence of a Borrelia burgdorferi polypeptide that may be
 CC useful in immunizing a subject against infection by a member of the
 CC Borrelia species.
 XX
 SQ Sequence 174 AA;

Query Match 100.0%; Score 875; DB 8; Length 174;
 Best Local Similarity 100.0%; Pred. No. 5.3e-79;
 Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QITQHILASILEESIIFLTRIFFAYIEDNDIFKKILOENKLYESSISFRFFYDENTKK 60
 DB 1 QITQHILASILEESIIFLTRIFFAYIEDNDIFKKILOENKLYESSISFRFFYDENTKK 60
 QY 61 KLEYKIIITIFNLLDKGSDAIFKFPVNGGLFSEDKVYLNNEGLLSISEIEILVKMLFF 120
 DB 61 KLEYKIIITIFNLLDKGSDAIFKFPVNGGLFSEDKVYLNNEGLLSISEIEILVKMLFF 120
 QY 121 EEKNIKDEKFKVYSRLDPKSGELYETLLEVDLRIADTTVHRIIEDGVYLIRTE 174
 DB 121 EEKNIKDEKFKVYSRLDPKSGELYETLLEVDLRIADTTVHRIIEDGVYLIRTE 174

RESULT 2
 ADU98757
 ID ADU98757 standard; protein; 1277 AA.
 XX
 AC ADU98757;
 XX
 DT 24-FEB-2005 (first entry)
 XX
 DE Borrelia burgdorferi antigenic polypeptide seqid 16.
 XX
 KW antibacterial; vaccine; immune stimulation; immunity; antigen;
 KW DNA library.
 XX
 OS Borrelia burgdorferi.
 XX
 PN WO2004103269-A2.
 XX
 PD 02-DEC-2004.
 XX
 PF 17-OCT-2003; 2003WO-US033056.
 XX
 PR 18-OCT-2002; 2002US-0419401P.
 XX
 PA (MACR-) MACROGENICS INC.
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX
 PI Sykes KF, Hale KS, Johnston SA;
 XX
 DR WPI; 2004-834155/82.
 DR N-PSDB; ADU98756.
 XX
 PT Immunizing a subject against Borrelia burgdorferi infection comprises
 PT providing to the subject at least one Borrelia antigen or its fragment.
 XX
 PS Claim 27; SEQ ID NO 16; 121pp; English.
 XX

CC The invention describes a method of immunizing a subject comprising
 CC providing to the subject at least one Borrelia antigen or its fragment to
 CC induce an immune response. Also described are: an isolated polynucleotide
 CC comprising a sequence having at least 17 contiguous nucleotides in common
 CC with a sequence not given in the specification; an isolated polypeptide
 CC having at least 5 consecutive amino acids of the sequence not given in
 CC the specification; a vaccine composition comprising at least one Borrelia
 CC antigen or at least one polynucleotide encoding a Borrelia antigen;
 CC screening for at least one test polypeptide or test polynucleotide
 CC encoding a polypeptide for an ability to produce an immune response;
 CC preparing a vaccine; vaccinating a subject; treating a subject infected

CC with a pathogen; passive immunization; vaccination; detecting Herpesvirus
 CC ; and detecting antibodies to a herpesvirus. The method is useful in
 CC immunizing a subject against Borrelia burgdorferi infection. This is the
 CC amino acids sequence of a Borrelia burgdorferi polypeptide that may be
 CC useful in immunizing a subject against infection by a member of the
 CC Borrelia species.
 XX
 SQ Sequence 1277 AA;

Query Match 99.7%; Score 872; DB 8; Length 1277;
 Best Local Similarity 99.4%; Pred. No. 1.5e-77;
 Matches 173; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QITQHILASILEESIIFLTRIFFAYIEDNDIFKKILOENKLYRSSISFRFFYDENTKK 60
 DB 268 EITQHILASILEESIIFLTRIFFAYIEDNDIFKKILOENKLYRSSISFRFFYDENTKK 327
 QY 61 KLEYKIIITIFNLLDKGSDAIFKFPVNGGLFSEDKVYLNNEGLLSISEIEILVKMLFF 120
 DB 328 KLEYKIIITIFNLLDKGSDAIFKFPVNGGLFSEDKVYLNNEGLLSISEIEILVKMLFF 387
 QY 121 EEKNIKDEKFKVYSRLDPKSGELYETLLEVDLRIADTTVHRIIEDGVYLIRTE 174
 DB 388 EEKNIKDEKFKVYSRLDPKSGELYETLLEVDLRIADTTVHRIIEDGVYLIRTE 441

RESULT 3
 AAU48711
 ID AAU48711 standard; protein; 1629 AA.
 XX
 AC AAU48711;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #9607.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US012865.
 XX
 PR 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI; 2001-616774/71.
 DR N-PSDB; AAS59543.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX
 PS Example 1; SEQ ID NO 9906; 1069pp; English.
 XX

CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central


```

CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1629 AA;

Query Match 12.5%; Score 109; DB 4; Length 1629;
Best Local Similarity 19.5%; Pred. No. 0.27;
Matches 36; Conservative 41; Mismatches 66; Indels 42; Gaps 6;

QY 13 ESIIFILRIFFIAYIEDN-----DIFKKILQENKLYRSSISPRYFF 53
Db 383 QSLRFLYRIILFLYAEASPELVLPVCAPOYDGSYLDRLRELVOVPLVTVQSRWAGTHL- 441
QY 54 YDENTKKKLYKKIITIFNLLDKGS-----DA----IKPPVPNGGLFSEDVKYKILNNEGLL 105
Db 442 -----YESLGTLFRLVDRGNLEDVDDATTGLSFHSLRADLFRPQATALIDVGL-L 490
QY 106 SISEIEILVKMLFFPE---KNIKDEFVKYKSLRDPKSFGEIYHTLLEYDLRIADTTVHR 162
Db 491 GNOALQOVLERLLLSKEQHGRKGRGYISYAEIGINQLGAVYEGMLSYSGFFATEDLYE 550
QY 163 IIEDG 167
Db 551 VANDG 555

RESULT 4
ID BM45230
XX ABM45230 standard; protein; 1629 AA.
AC ABM45230;
XX
XX 20-OCT-2003 (first entry)
XX
XX Propionibacterium acnes predicted ORF-encoded polypeptide #9906.
XX
XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX immunostimulant; immune response; vaccine.
XX
XX Propionibacterium acnes.
XX
XX W02003033515-A1.
XX
XX 24-APR-2003.
XX
XX 11-OCT-2002; 2002WO-US032727.
XX
XX 15-OCT-2001; 2001US-00978825.
XX
XX (CORI-) CORIXA CORP.
XX
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
XX Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
XX Barth B, Vallieue-Douglas J;
XX
XX WPI; 2003-381789/36.
XX
XX N-PSDB; ACF64472.
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
XX polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
XX or for stimulating an immune response specific for a P. acnes protein.

```


XX 22-MAR-2002; 2002WO-US008992.
PF
XX 23-MAR-2001; 2001US-00816703.
PR 17-OCT-2001; 2001US-00983000.
XX
XX (AGYT-) AGY THERAPEUTICS INC.
XX
XX Mueller S, Melcher T, Chin DJ;
XX
XX WPI; 2003-029903/02.
DR N-PSDB; AAD48126.
XX
XX Developing active agents that modulate the activity of a brain tumor
PT protein target gene or gene product for treating e.g. stroke or cancer,
PT comprises contacting an agent with a brain tumor protein.
XX
XX Disclosure; Page 84-85; 135pp; English.
XX
XX The invention relates to a method for developing biologically active
CC agents that modulate activity of a brain tumor protein target (Tbt) gene
CC or gene product. The method is useful for developing biologically active
CC agents that modulate the activity of a brain tumor protein target gene
CC or gene product. Compounds that bind to the brain tumor proteins are
CC useful for treating e.g. ischaemic stroke, brain cancer, epilepsy,
CC schizophrenia, depression, Alzheimer's disease, Parkinson's disease,
CC Huntington's chorea, traumatic head injury, dementia, stupor, headache,
CC coma, vertigo, weakness, myasthenia gravis, cerebrovascular disorders,
CC infectious disorders (including fungal, bacterial, viral and parasitic
CC infections), multiple sclerosis, and other complications associated with
CC pregnancy, medical illness, alcohol and substance abuse, toxins and
CC metabolic deficiencies. The brain tumor proteins may also be used to
CC raise antibodies. The present sequence is human protein tyrosine
CC phosphatase zeta SM1 variant protein used to illustrate the method of the
CC invention
XX
SQ Sequence 374 AA;

Query Match 11.1%; Score 97; DB 6; Length 374;
Best Local Similarity 21.8%; Pred. No. 0.6;
Matches 48; Conservative 36; Mismatches 50; Indels 86; Gaps 11;

QY 26 YIENDIF---KKIQENKLYRSSIPRYFFYDENTKKLEYKII----- 68
Db 156 YCFDADRSSPEEAVKGGKRALSLF-----EVGTEENLDFKAIIDGVESVRFKQA 210
QY 69 -----TIFNLDKGSDAIKPFVFNGLFS---EDKVYLNNEGLLSIEIB-----EIL- 114
Db 211 ALDPFILLNLLPNSTD--KYIYNGSLTSPCTDVTVDWIVFKDTVSISESQAVFCEVLT 268
QY 115 -----VKMLFEEKNIKDEKFKYSR-----LDPKSFGE 143
Db 269 MQQSGYMLMDYLNQNFREQY-KFSRQVFSYTGKEIHAVCSSPENQVADPENYTS 327
QY 144 L-----YETLLE-----YDLRIADTVVHRIEDG 167
Db 328 LLVTWERPRVVDTMIEKFAVLYQQLDGEDQTKHEFLTDG 367

RESULT 10
ADL23826
ID ADL23826 standard; protein; 389 AA.
XX
AC ADL23826;
XX
DT 03-JUN-2004 (first entry)
XX
XX Pyrococcus furiosus intein Pfu IF2 amino acid sequence.
XX
XX cyclic peptide; interacting protein inhibition; fusion nucleic acid;
KW intein motif; intein; intein-catalysed cyclic peptide library;
KW cancer cell line; cardiovascular; neurobiology; bone biology;
KW biotechnology; skin biology; endocrinology; drug toxicity;

KW drug resistance; drug development.
XX
OS Pyrococcus furiosus.
XX
PN WO2004019890-A2.
XX
PD 11-MAR-2004.
XX
XX 30-AUG-2003; 2003WO-US027371.
XX
XX 30-AUG-2002; 2002US-00232758.
PR 23-APR-2003; 2003US-00422536.
XX
XX (RIGS-) RIGEL PHARM INC.
XX
XX Lorens JB, Pray T, Kinsella TM, Bennett MK;
PI WPI; 2004-239115/22.
XX
XX Generating intracellular cyclic peptide libraries, useful in identifying
PT interaction of interacting proteins by providing cells comprising first
PT and second interacting protein, and contacting cells with library of
PT fusion nucleic acids.
XX
XX Disclosure; Fig 3R; 156pp; English.
XX
XX The present invention describes a method (M1) for identifying a cell
CC comprising cyclic peptides which prevent interaction of interacting
CC proteins. M1 comprises providing cells with first and second interacting
CC proteins and contacting cells with a library of fusion nucleic acids
CC comprising (from 5' to 3'): a nucleic acid encoding a C-terminal intein
CC motif, a peptide, and an N-terminal intein motif, under conditions such
CC that cyclic peptide is formed. Also described is a method (M2) for
CC isolating the nucleic acid that encodes the cyclic peptide which when
CC introduced into cell caused the change in ubiquitination of the molecule,
CC comprising contacting a cell culture with an intein-catalysed cyclic
CC peptide library such that cells in the culture receive not more than one
CC member of the library, and monitoring the cells for a change in
CC ubiquitination of a molecule. M1 is useful for identifying a cell
CC comprising a cyclic peptide which is capable of preventing interaction of
CC interacting proteins. M1 is useful in creating cell lines from cancer
CC patients and in cardiovascular, neurobiology, drug toxicities and drug
CC biotechnology, skin biology, endocrinology, drug toxicities and drug
CC resistance applications. M1 is useful for identifying agents that
CC regulate intracellular and sarcolemmal calcium cycling in
CC cardiomyocytes, agents that diminish embolic phenomena in arteries and
CC arterioles, and for screening peptides that block transport mechanisms.
CC M1 is useful for screening enhanced contractile properties of
CC cardiomyocytes. M1 enables screening of a large number of cell types
CC under a wide variety of conditions. M1 enables improvement in the
CC performance of existing or developmental drugs. The present sequence is
CC used in the exemplification of the present invention.
XX
SQ Sequence 389 AA;

Query Match 11.1%; Score 97; DB 8; Length 389;
Best Local Similarity 20.6%; Pred. No. 0.64;
Matches 45; Conservative 38; Mismatches 79; Indels 56; Gaps 7;

QY 5 HILASILEESIIFILRIFFIAYIEDNDIFKKIQENKLYRSSISF----- 49
Db 179 YIEVPLFEDPLLLIRFGIVSRIEKSTLKISGKNLELFRKHVGTDSKAKALDELISK 238
QY 50 -----RYFFYDE-----NTKKLE-----YKKIITIFNLDKGSD--AIKFPYFN 87
Db 239 AKESERYPIIEELRRLGLLFGFTRENLRIEENPTYEIVMEILERIERGSPNLAEKIAYLE 298
QY 88 GGLFSEDKVKYLNNEGLLSISEIEILVKMLFPEEKNIKDEKFKYSRLDPKSFGEIYET 147
Db 299 GRIKEENYLRILEEGEGLTENGKLTGKLELVWNRNREFDSKQVDYVR-----NIVEN 351
QY 148 L-----LEYDLRIADTV--VHRIEDGVYLIET 173

Db 352 LVFLPVEKVERIEYEGYVDVTTETHNFVANGILVHNT 389

RESULT 11
ADQ18134
ID ADQ18134 standard; protein; 1005 AA.

AC ADQ18134;

DT 26-AUG-2004 (first entry)

DE Human soft tissue sarcoma-upregulated protein - SEQ ID 951.

XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.

KW Homo sapiens.

OS WO2004048938-A2.

FN 10-JUN-2004.

PD 26-NOV-2003; 2003WO-US038193.

PF 26-NOV-2002; 2002US-0429739P.

XX (PROT-) PROTEIN DESIGN LABS INC.

XX Aziz N, Ginsburg WM, Zlotnik A;

XX WPI; 2004-441208/41.

DR Early detection of soft tissue sarcoma comprises determining expression of a gene in a first soft tissue sample and a normal soft tissue sample and comparing the gene expression, also useful in treating soft tissue sarcoma.

XX Example 2; SEQ ID NO 951; 210pp; English.

XX The invention relates to a novel method for detecting soft tissue sarcoma which comprises obtaining a first soft tissue sample from an individual and a normal soft tissue sample from the same or different individual, determining the expression of a gene in both samples and comparing the expression of the gene in both soft tissue samples, where a higher level of protein expression in the first soft tissue sample indicates the presence of soft tissue sarcoma. The method of the invention has cytostatic applications and may be useful for detecting soft tissue sarcoma, possibly via gene therapy or vaccine production. The nucleic acid sequences may be useful in diagnostic and screening applications. The current sequence is that of a human soft tissue sarcoma-upregulated protein of the invention. The current sequence is not shown within the specification per se but was submitted in CD format by the inventor.

XX Sequence 1005 AA;

Query Match 11.1%; Score 97; DB 8; Length 1005;
Best Local Similarity 21.8%; Pred. No. 2.2;
Matches 48; Conservative 36; Mismatches 50; Indels 86; Gaps 11;

QY 26 YIEDNDIF----KKTLQNKLYRISIPRYFYDNTKKLEYKKII----- 68

Db 156 YCFDADRFSFEAVKGGKRLALSLP-----EVGTENLDFFKALIDGVESVRFKQA 210

QY 69 -----TIFNLDKGSDAKFFVFNGLFS----EDKVKYLNNEGLLSISEIE-----EIL- 114

Db 211 ALDPPILLNLLPNSD--KYIYNGSLTSPPTDVTVDWIVFKDTVISQSILAVFCEVLT 268

QY 115 -----VKMLPFEENKIKDEKPKVKSR-----LDPKSGE 143

Db 269 MQQSGYVLMYLDLQNNFREQQY-KFSRQVFSYTGKEIHEAACRSEPNVQADPENYTS 327

QY 144 L-----YETLLE-----YDLRIADTVVHRIEDG 167

Db 328 LLVTWTPRVVYDTMIKFAVLYQQLDGEDQTKHEFLTDG 367

RESULT 12
ADN39946

ID ADN39946 standard; protein; 1113 AA.

AC ADN39946;

DT 17-JUN-2004 (first entry)

DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C316.

XX Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
KW vulnerary; gene therapy; vaccine.

OS Homo sapiens.

FN WO2003042661-A2.

XX 22-MAY-2003.

XX 13-NOV-2002; 2002WO-US036810.

XX 13-NOV-2001; 2001US-0350666P.

XX 21-NOV-2001; 2001US-0332464P.

XX 29-NOV-2001; 2001US-0334393P.

XX 03-DEC-2001; 2001US-0335394P.

XX 14-DEC-2001; 2001US-0340376P.

XX 08-JAN-2002; 2002US-0347211P.

XX 10-JAN-2002; 2002US-0347349P.

XX 08-FEB-2002; 2002US-0355250P.

XX 13-FEB-2002; 2002US-0356714P.

XX 20-FEB-2002; 2002US-0359077P.

XX 29-MAR-2002; 2002US-0368809P.

XX 04-APR-2002; 2002US-0370110P.

XX 12-APR-2002; 2002US-0372246P.

XX 05-JUN-2002; 2002US-0386614P.

XX 16-JUL-2002; 2002US-0396839P.

XX 22-JUL-2002; 2002US-0397775P.

XX 22-JUL-2002; 2002US-0397845P.

XX 09-SEP-2002; 2002US-0409450P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;

XX Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;

XX WPI; 2003-468649/44.

XX N-PSDB; ADN39729.

XX Determining the presence or absence of a pathological cell in a patient,

XX useful for diagnosing, prognosing or treating cancer, comprises detecting

XX a nucleic acid in a biological sample.

XX Claim 12; SEQ ID NO C316; 1385pp; English.

XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064) whose expression is upregulated or downregulated in specific cancers or other diseases such as angiogenic or fibrotic disorders, and to methods of determining the presence or absence of a pathological cell in a patient by detecting a nucleic acid at least 80% identical to those of the invention or by detecting a polypeptide of the invention. The invention also relates to expression vectors and host cells comprising a nucleic acid of the invention; antibodies which specifically bind a polypeptide of the invention; use of such antibodies for drug targeting; and methods of screening for modulators of activity or expression of the polypeptides and nucleic acids. The nucleic acids, polypeptides, and antibodies and methods are useful for diagnosing, prognosing and treating

CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.
XX
SQ Sequence 1113 AA;
Query Match 11.1%; Score 97; DB 7; Length 1113;
Best Local Similarity 21.8%; Pred. No. 2.5;
Matches 48; Conservative 36; Mismatches 50; Indels 86; Gaps 11;
Qy 26 YIEDNIF----KKILOENKLYRSSISFRYPFFYDENTKKLEYKKII----- 68
Db 156 YCFDADRFSSEAEVKGKGLRALSILF-----EVGTEENLDKFAIIDGVESVSRFGKQA 210
Qy 69 -----TIFNLLDKGSDAIKPPVFNGLFS-----EDKVKYLNNEGLLSISEIE-----BIL- 114
Db 211 ALDPPFILLNLLPNSTD--KYIYNGSLTSPPTCTDWDIVFKOTVSISESQLAVFCEVLT 268
Qy 115 -----VKMLFFEEKNIKDEKFKVYSR-----LDPKSFG 143
Db 269 MQQSGYVMLDYLNQNNFREQQY-KFSRQVFSYTGKEEIEHAVCSSEPNVQADPENYTS 327
Qy 144 L-----YETILE-----YDLRIADTTVHRIEDG 167
Db 328 LLVTWPRVYVDTMIEKFAVLYQQLDGEDQTKHEFLTGD 367
RESULT 13
ID ABU56496 standard; protein; 1337 AA.
XX AC ABU56496;
XX DT 02-APR-2003 (first entry)
XX DE Lung cancer-associated polypeptide #89.
XX KW Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX OS Unidentified.
XX FN WO200286443-A2.
XX PD 31-OCT-2002.
XX PF 18-APR-2002; 2002WO-US012476.
XX PR 18-APR-2001; 2001US-0284770P.
PR 10-MAY-2001; 2001US-0290492P.
PR 09-NOV-2001; 2001US-0339245P.
PR 13-NOV-2001; 2001US-0350666P.
PR 29-NOV-2001; 2001US-0334370P.
PR 12-APR-2002; 2002US-0372246P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Aziz N, Murray R;
XX DR WPI; 2003-093161/08.
XX DR N-PSDB; ABX76220.
XX PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
XX expression in lung cancer.
XX PS Claim 27; Page 262; 453pp; English.

XX The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridises
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful
CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABU56408-ABU56745 represent lung cancer-associated polypeptides of the
CC invention
XX
SQ Sequence 1337 AA;
Query Match 11.1%; Score 97; DB 6; Length 1337;
Best Local Similarity 21.8%; Pred. No. 3.2;
Matches 48; Conservative 36; Mismatches 50; Indels 86; Gaps 11;
Qy 26 YIEDNIF----KKILOENKLYRSSISFRYPFFYDENTKKLEYKKII----- 68
Db 38 YCFDADRFSSEAEVKGKGLRALSILF-----EVGTEENLDKFAIIDGVESVSRFGKQA 92
Qy 69 -----TIFNLLDKGSDAIKPPVFNGLFS-----EDKVKYLNNEGLLSISEIE-----BIL- 114
Db 93 ALDPPFILLNLLPNSTD--KYIYNGSLTSPPTCTDWDIVFKOTVSISESQLAVFCEVLT 150
Qy 115 -----VKMLFFEEKNIKDEKFKVYSR-----LDPKSFG 143
Db 151 MQQSGYVMLDYLNQNNFREQQY-KFSRQVFSYTGKEEIEHAVCSSEPNVQADPENYTS 209
Qy 144 L-----YETILE-----YDLRIADTTVHRIEDG 167
Db 210 LLVTWPRVYVDTMIEKFAVLYQQLDGEDQTKHEFLTGD 249
RESULT 14
ID ABU56690 standard; protein; 1337 AA.
XX AC ABU56690;
XX DT 02-APR-2003 (first entry)
XX DE Lung cancer-associated polypeptide #283.
XX KW Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX OS Unidentified.
XX FN WO200286443-A2.
XX PD 31-OCT-2002.
XX PF 18-APR-2002; 2002WO-US012476.
XX PR 18-APR-2001; 2001US-0284770P.
PR 10-MAY-2001; 2001US-0290492P.
PR 09-NOV-2001; 2001US-0339245P.
PR 13-NOV-2001; 2001US-0350666P.
PR 29-NOV-2001; 2001US-0334370P.

Qy 69 -----TIFNLLDKGSDAIFPVFNGGLFS---EDKVKYLNNEGGLLSISEIE-----EIL- 114
Db 93 ALDPFILLNLLPNSTD--KYIYNGSLTSPCTDTVDWIVFKDTSISESOLAVFCEVLT 150
Qy 115 -----VKMLFFBEKNIKDEKFKYKSR-----LDPKSFG 143
Db 151 MQQSGYVMLMDYLQNNFREQQY-KFSRQVFSYTGKEIHEAVCSSEPNVQADPENYTS 209
Qy 144 L-----YETILE-----YDLRIADTTVHRIIDG 167
Db 210 LLVTWERPRVYDTMIEKFAVLYQQLDGEDQTKHEFLTGD 249

Search completed: January 24, 2006, 19:44:35
Job time : 40.8844 secs

RESULT 2
B70236
hypothetical protein BBH09 - Lyme disease spirochete plasmid H/lp28-3
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C:Accession: B70236
B70236
Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra,
son, D.; Peterson, J.; Kertavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; H
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:9805943; PMID:9403685

A:Accession: B70236
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1278 <KLE>
A:Cross-references: UNIPROT:O50667; UNIPARC:UPI0000056814; GB:AE000784; NID:g2690041; P
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid

Query Match 95.3%; Score 834; DB 2; Length 1278;
Best Local Similarity 92.5%; Pred. No. 1.8e-54;
Matches 161; Conservative 11; Mismatches 2; Indels 0; Gaps 0;

QY 1 QITQHILASILESIIFILRIFFIAYIEDNDIFKKILOENKLYRSSISPRYPYDENTKK 60
DB 274 EITRHILASILESIIFILRIFFIAYIEDNDIFKKILEENKLYRSSVSFRYPYDENTKK 333

QY 61 KLEYKKIITIFNLLDKGSDAIFKPFVNGGLFSEDKVYLNNEGILLSIEIEILVKMLFF 120
DB 334 KLEYKKIITIFNLLDKGSDAIFKPFVNGGLFAQDKVYLNNEGILLSIEIEILVKILFF 393

QY 121 EERNKIDKFKVYSLRDPKSGFGLYETLLYDRIADTTVHRIEDGVILIRTE 174
DB 394 EERNKIDKFKVYSLRDPKSGFGLYETLLYDRIADTTVHRIEDGVILIRTE 447

RESULT 3
E81339
probable restriction/modification enzyme Cj0690c [imported] - Campylobacter jejuni (str
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: E81339
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barr
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals h
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: E81339
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1250 <PAR>
A:Cross-references: UNIPROT:Q9PPL7; UNIPARC:UPI00000CICEE; GB:AL1139076; GB:AL111168; NID
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj0690c

Query Match 26.5%; Score 232; DB 2; Length 1250;
Best Local Similarity 33.7%; Pred. No. 1.4e-09;
Matches 55; Conservative 34; Mismatches 68; Indels 6; Gaps 3;

QY 3 TQHTLASILESIIFILRIFFIAYIEDNDIFKKILOENKLYRSSISPRYPF--YDENTKK 60
DB 260 TKADLKLIYENSLYIFPRLFIAYFEDK--FEIILEKHKFKSKISLTILENLQDESS 317

QY 61 KLEYKKIITIFNLLDKGSDAIFKPFVNGGLFSEDKVYLNNEGILLSIEIEILVKMLFF 120
DB 318 SGGEGELENFIYNYKKGFMDFVFNGLFDESKTALLSTPKFNDKDKLFILNQLNF 377

QY 121 EERNKIDKFKVYSLRDPKSGFGLYETLLYDRIADTTVHRI 163
DB 378 KDKNLSFKR--DYKTLVSEHLGTIVEGLLSVFFPEIANEDIVYV 418

RESULT 4
T08313
conserved hypothetical protein H1130 - Halobacterium sp. (strain NRC-1) plasmid pNRC100
C:Species: Halobacterium sp.
A:Variety: strain NRC-1
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T08313
R:Ng, W.V.; Clufo, S.A.; Smith, T.M.; Bumgarner, R.E.; Baekin, D.; Faust, J.; Hall, B.;
Genome Res. 8, 1131-1141, 1998
A:Title: Snapshot of a large dynamic replicon in a halophilic Archaeon: megaplasmid or m

A:Reference number: Z16408; MUID:99063795; PMID:9847077
A:Accession: T08313
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-706 <DAS>
A:Cross-references: UNIPROT:O52000; UNIPARC:UPI00000631E7; EMBL:AF016485; NID:g28222278;
A:Experimental source: strain NRC-1
C:Genetics:
A:Gene: HALOSP:H1130
A:Genome: plasmid pNRC100

Query Match 20.9%; Score 182.5; DB 2; Length 706;
Best Local Similarity 27.6%; Pred. No. 3.6e-06;
Matches 55; Conservative 39; Mismatches 58; Indels 47; Gaps 10;

QY 10 ILEESIIFILRIFFIAYIEDNDIFKKILO-ENKLYRSSISF-----RYF 52
DB 274 IHDSLIYLRILFVLYAEAG--RELLDTSIEIYEQSYSLNSLKGIEIAEELDSGDPKVR 331

QY 53 FYDENTKKLEYKKIITIFNLLDKGS-----DAIKFPVFNGLFSED-----KVKY 98
DB 332 DQDNLOSLD-----ELFTLINKSKSGRIGIEEDLYIPAYNGELFRTDPDEDDSAEAKF 386

QY 99 INNE--GLLSISIEIEILVKMLFFBEKNIKDEK-FVKYSLRDPKSGFGLYETLLYDLRI 155
DB 387 LANHDVGDAYLAKVWELLTR----SKNDGGGKIFVDYSYSLDVRHLGSIYEGLEQLNV 442

QY 156 ADTTHRIEDGVILIRTE 174
DB 443 ADEPL--ALDDGEYVSADE 459

RESULT 5
F91191
probable LPS biosynthesis enzyme waaJ [imported] - Escherichia coli (strain O157:H7, su
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: F91191
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Iehii, K.; Yokoyama, K.; Han, C.G
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Experimental source: strain O157:H7, substrain RIMD 050952
C:Genetics:
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <HAY>
A:Cross-references: UNIPROT:Q8XDC2; UNIPARC:UPI0000000396; GB:BA000007; PIDN:BA837925.1
A:Experimental source: strain O157:H7, substrain RIMD 050952
C:Genetics:

C:Superfamily: lipopolysaccharide glucosyltransferase I

Query Match 11.9%; Score 104.5; DB 2; Length 337;
Best Local Similarity 21.0%; Pred. No. 1.1;
Matches 45; Conservative 46; Mismatches 74; Indels 49; Gaps 8;

QY 1 QITQHILASILESIIFIL-----RIFIAIYENDIDFKILOEN----KLYRSSISF 49
DB 57 KFTFHIADYLDKEYIELLSQLATKYQTVIKLYHIDSEPLKALPQSNIPVSIYRLLSF 116

QY 50 RYF-----FYDENTKKK-----LEYKK-----IITIFNLLDKGSDAIFKPFVNG 88
DB 117 DYFSARLDLSLYLDADIVCKGSLNELIALEFKDEYGNVIVDMQSKSAERLCNEDFNG 176

QY 89 GLFSEDKVYLNNE-----GLLSISIEIEILVKMLFFBEKNIKDEKFKVYKYSRL 136
DB 177 SYFNSG-VMYNLRWMLKQRLTEKFFDLS-----DESIKKLYKYPQDILNMLFHHAKI 231

QY 137 DPKSFGELYETLLYDRIADTTVHRIEDGVIL 170
DB 232 LPRKYNCIYTIKSEFEKSEKNSEYTRFINDDTVFI 265

A;Residues: 1479-2091 <KR2>
A;Cross-references: UNIPARC:UPI000016AP42; GB:X54135; NID:G35795; PIDN:CAA38070.1; PID:G
C;Genetics:
A;Gene: GDB:PTPR21; PTP2; PTP2; PTP2; PTP18; RPTPB
A;Cross-references: GDB:127353; OMIM:176891
A;Map position: 7q31.3-7q31.3
C;Function:
A;Description: catalyzes the hydrolysis of peptidyl-phosphotyrosine to release phosphate
A;Note: may be involved in the regulation of specific developmental processes in the cen
C;Superfamily: protein-tyrosine-phosphatase, receptor type beta; carbonic anhydrase hom
C;Keywords: alternative splicing; brain; glycoprotein; phosphoprotein; phosphoric monoo
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-24/Domain: protein-tyrosine-phosphatase, receptor type zeta #status predicted <M
F;25-1635/Domain: extracellular #status predicted <EXT>
F;25-754,1614-2314/Product: protein-tyrosine phosphatase receptor type zeta, short form
F;38-300/Domain: carbonic anhydrase homology <CAH>
F;1636-1661/Domain: transmembrane #status predicted <TMN>
F;1662-2314/Domain: intracellular #status predicted <INT>
F;1749-1980/Domain: protein-tyrosine-phosphatase homology <PTP1>
F;2047-2270/Domain: protein-tyrosine-phosphatase homology <PTP2>
F;105,134,223,232,324,381,480,497,501,552,602,629,677,782,1017,1050,1082,1122,1456,1561,
F;1932/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1938/Binding site: substrate phosphate (Arg) #status predicted

Query Match 11.1%; Score 97; DB 1; Length 2314;
Best Local Similarity 21.8%; Pred. No. 31;
Matches 48; Conservative 36; Mismatches 50; Indels 86; Gaps 11;

QY 26 YIEDNDIF----KKILQENKLYRSSISFRFFYDENTKKLYKKII----- 68
Db 156 YCFDADRFSSFEAVKGGKRLALSILF-----EVCTEENLDFKALIDGVESVSRCKQA 210

QY 69 -----TIFNLLDKGDAIKFPVNGGLFS---EDKVKYLNNEGILLSISEI---EIL- 114
Db 211 ALDPFILLNLLNSTD---KYYIYNGSLTSPCTDVTDMWIFKDTVSISEQLAVFCEVLT 268

QY 115 -----VKMLPFEEKIKDEKFKVYSR-----LDPKSFG 143
Db 269 MQQSGVYMLDYLQNNFREQY---KFSRQVSSVTGKEETHEAVCSSEPNVQADPENYTS 327

QY 144 L-----YETLLE-----YDLRIADTTVHRIEDG 167
Db 328 LLVTWERPVVYDTMTEKFAVLVQQLDGEDQTKHEFLTDG 367

RESULT 13
T18414
protein g377 - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18414
R;Handman, E.; Osborn, A.H.; Symons, F.; Van Driel, R.; Cappai, R.
Mol. Biochem. Parasitol. 74, 143-156, 1995
A;Title: The Leishmania promastigote surface antigen 2 complex is differentially expres
A;Reference number: Z18933; MUID:96360472; PMID:8719156
A;Accession: T18414
A;Status: preliminary; translated from GB/EMBL/DDBU
A;Molecule type: DNA
A;Residues: 1-3119 <HAN>
A;Cross-references: UNIPROT:Q25857; UNIPARC:UPI0000082072; EMBL:L04161; NID:G309687; PID

Query Match 11.1%; Score 97; DB 2; Length 3119;
Best Local Similarity 25.7%; Pred. No. 43;
Matches 38; Conservative 23; Mismatches 41; Indels 46; Gaps 7;

QY 18 ILRIPIAYIEDNDIPKKILQENKLYRSSISFRFFYDENTKKLYKKIIITFNLLDKG 77
Db 2599 IIEPTTKFIENKPVFKNM---PLLXNRIKY-----DSINEKLOYKEELDLDHMDNE 2649

QY 78 SDAIKFPVNGGLFSDEKVKYLNNEGILLSISEIILVKMLFF-----EEKIKD--- 127
Db 2650 S-----FHQ-----LLDLINKESITLQSLFSPRYTEHANIKNDFPK 2687

QY 128 -----EKFKVYSRLDPKPSFGELYETILLEY 151
Db 2688 LPMFKKXYTVVQ--DEKFFINLYEYLYNY 2714

RESULT 14
B64516
hypothetical protein MJEC502 - Methanococcus jannaschii plasmid PURB801
C;Species: Methanococcus jannaschii
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: B64516
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
; Reich, C.J.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese,
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschi
A;Reference number: A64300; MUID:96337999; PMID:8688087
A;Accession: B64516
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1181 <BUL>
A;Cross-references: UNIPROT:Q60301; UNIPARC:UPI000013C258; GB:L77119; NID:G1500688; TIG
C;Genetics:
A;Map position: ECSREV4814-1269
A;Genome: plasmid
A;Start codon: GTG
A;Note: this stable 16-kilobase pair plasmid is also designated ECS (small extrachromos

Query Match 11.0%; Score 96.5; DB 2; Length 1181;
Best Local Similarity 22.4%; Pred. No. 16;
Matches 36; Conservative 32; Mismatches 46; Indels 47; Gaps 8;

QY 14 SIIFILRIPIAYIED-----NDIFKKILQENKLYRSSISFRYF-----FYD----- 55
Db 290 AVLLMRLIFIKLEDKGIVPRDLRLTYEDYK--KSNVLINVDAYLPLFVEVLNTP 347

QY 56 -----ENTKKLYKKIIITFNLLDKGDAIKFPVNGGLFSDEKVKYLNNEGILLSISEI 111
Db 348 DERKENIRTNPYKDI-----PYLNGGLFRSNVNP---NELSFTIKD-N 387

QY 112 EILVKMLFEEK---NIKDEKFKVYSRLDPKPSFGELYETIL 149
Db 388 EIIGVINFLERYKFTLSTSEGSEVELNPDLIGYVYEKLI 428

RESULT 15
S38614
hypothetical protein 558 (tufa 3' region) - euglenid (Astasia longa) plastid
C;Species: plastid Astasia longa
C;Date: 31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change 09-Jul-2004
C;Accession: S38614
R;Gockel, G.; Baier, S.; Hachtel, W.
submitted to the EMBL Data Library, November 1993
A;Reference number: S38590
A;Accession: S38614
A;Molecule type: DNA
A;Residues: 1-558 <GOC>
A;Cross-references: UNIPROT:P34782; UNIPARC:UPI000013A77B; EMBL:X75652; NID:G414919; PI
C;Keywords: chloroplast; plastid

Query Match 10.9%; Score 95; DB 2; Length 558;
Best Local Similarity 23.2%; Pred. No. 9.3;
Matches 39; Conservative 34; Mismatches 59; Indels 36; Gaps 7;

QY 22 FTIAYIEDNDIPKKILQENKLYRSSISFRFFYDENTKKLYKKIIITFNLLDK--GSD 79
Db 72 YFTINLEDFKTIKFTDTSVLIVSVGTRFFHSESEQKDS-----LIFNFIPIEINS 126

QY 80 AIKFPVNGGLFSDEKVKY-----LNNEGILLSISEIILVKMLFF 120
Db 127 GL---MGTKVLLESKKILYENENKEKDLKLVHIFPDIINNVLIGVKVLE---KKKIFY 181

Qy 121 EEKNKDEKFKYSRLDPKSF--GELYETLLEYDLRIADTTVHRIED 166
Db 182 ENENIKOLEFIEGSELEGGQIKSEDIEEKVVDKY---DDSEGRVLLD 226

Search completed: January 24, 2006, 19:59:06
Job time : 8.60107 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:27:53 ; Search time 37.3889 Seconds
(without alignments)
3283.383 Million cell updates/sec

Title: US-10-688-058-14
Perfect score: 875
Sequence: 1 QITQHILASILESIIFILR.....IADTVHRIEDGVYLIRTE 174

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_05-80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	872	99.7	1277	2	O50698 BORBU	O50698 borrelia bu
2	834	95.3	1278	2	O50667 BORBU	O50667 borrelia bu
3	821	93.8	1194	2	O5XYT2_BORGA	O5XYT2 borrelia ga
4	470	53.7	149	2	O5XYI6_BORGA	O5XYI6 borrelia ga
5	290	33.1	806	2	O5XYG5_BORGA	O5XYG5 borrelia ga
6	232	26.5	1250	2	O5HV91_CAMJUR	O5HV91 campylobact
7	232	26.5	1250	2	O5PPL7_CAMJUR	O5PPL7 campylobact
8	182.5	20.9	662	2	O9HI06_HALSA	O9HI06 halobacteri
9	182.5	20.9	706	2	O52000_HALSA	O52000 halobacteri
10	109.5	12.5	1193	2	O50N37_ENTHI	O50N37 entamoeba h
11	109.5	12.5	2181	2	O51017_ENTHI	O51017 entamoeba h
12	108.5	12.4	2337	2	O4XXL0_PLACH	O4XXL0 plasmodium
13	107.5	12.3	868	2	O54HE3_DICDI	O54HE3 dictyosteli
14	107.5	12.3	939	2	O5FM50_LACAC	O5FM50 lactobacilli
15	106.5	12.2	1524	2	O8FSZ4_COREF	O8FSZ4 corynebacte
16	106	12.1	1572	2	O4XYM9_PLACH	O4XYM9 plasmodium
17	105.5	12.1	309	2	O834E5_ENTFA	O834E5 enterococcu
18	105.5	12.1	837	2	O6F1B2_MESFL	O6F1B2 mesoplasma
19	105	12.0	634	2	O7RS91_PLAYO	O7RS91 plasmodium
20	104.5	11.9	337	2	O83PN9_SHIFL	O83PN9 shigella fl
21	104.5	11.9	337	2	O8XDC2_ECOS7	O8XDC2 escherichia
22	104.5	11.9	867	2	O9PJ80_CAMJUR	O9PJ80 campylobact
23	104.5	11.9	878	2	O4HEH3_CAMCO	O4HEH3 campylobact
24	104.5	11.9	1084	2	O9PRL8_UREPA	O9PRL8 ureaplasma
25	104.5	11.9	1257	2	O5HXK7_CAMJUR	O5HXK7 campylobact
26	103	11.8	687	2	O4XNV9_PLACH	O4XNV9 plasmodium
27	102.5	11.7	788	2	O7RIG0_PLAYO	O7RIG0 plasmodium
28	102	11.7	1843	2	O8IC40_PLAF7	O8IC40 plasmodium
29	101.5	11.6	609	2	O9P193_CAMJUR	O9P193 campylobact
30	101.5	11.6	2259	1	YCF2_PHYPA	P61243 physcomitre
31	100.5	11.5	723	2	O8D314_WIGBR	O8D314 wigglewort

32	100.5	11.5	970	2	O813G3_PLAF7	O813G3 plasmodium
33	100.5	11.5	1610	2	O5FQ08_GLUOX	O5FQ08 gluconobact
34	100	11.4	856	2	O4Z0W0_PLABE	O4Z0W0 plasmodium
35	100	11.4	1056	2	O7VII1_HELHP	O7VII1 helicobacte
36	99.5	11.4	258	2	O812E9_PLAF7	O812E9 plasmodium
37	99.5	11.4	371	2	O4FNN9_9RICK	O4FNN9 candidatus
38	99.5	11.4	486	2	O4ZIT6_PLABE	O4ZIT6 plasmodium
39	99.5	11.4	3377	2	O81EH5_PLAF7	O81EH5 plasmodium
40	99	11.3	273	2	O97GM0_CLOAB	O97GM0 clostridium
41	99	11.3	290	2	O7NAL6_MYCGA	O7NAL6 mycoplasma
42	99	11.3	661	1	PALA_YEAST	O12033 saccharomyc
43	99	11.3	930	2	O6MAP0_PARUM	O6MAP0 parachlamyd
44	98	11.2	714	2	O4YHMS_PLABE	O4YHMS plasmodium
45	97.5	11.1	291	2	O4MUDI_BACE	O4MUDI bacillus ce

ALIGNMENTS

RESULT 1

ID	O50698 BORBU	PRELIMINARY;	PRT; 1277 AA.
AC	O50698 BORBU	PRELIMINARY;	PRT; 1277 AA.
DT	01-JUN-1998	(TREMBlrel. 06, Created)	
DT	01-JUN-1998	(TREMBlrel. 06, Last sequence update)	
DT	01-MAR-2004	(TREMBlrel. 26, Last annotation update)	
DE	Hypothetical protein BBEO2.		
GN	OrderedLocusNames=BBEO2;		
OS	Borrelia burgdorferi (lyme disease spirochete).		
OG	Plasmid lp25.		
OC	Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;		
OC	Borrelia burgdorferi group.		
OX	NCBI_TaxID=139;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=ATCC 35210 / B31;		
RX	MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;		
RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,		
RA	Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,		
RA	Gwin M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,		
RA	Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,		
RA	Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,		
RA	Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A.,		
RA	Artiach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,		
RA	Roberts K.M., Hatch B., Smith H.O., Venter J.C.;		
RT	"Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi."		
RL	Nature 390:580-586(1997).		
DR	EMBL; AE000785; AAC66031.1; -; Genomic_DNA.		
DR	PIR; E70224; E70224.		
DR	HSSP; Q57597; IMJH.		
DR	TIGR; BBEO2; -		
DR	GO; GO:0003677; F:DNA binding; IEA.		
DR	GO; GO:0008170; F:N-methyltransferase activity; IEA.		
DR	GO; GO:0006306; P:DNA methylation; IEA.		
DR	InterPro; IPR007409; DUF450.		
DR	InterPro; IPR002296; N12N6_mtfase.		
DR	Pfam; PF04313; HSDR N; 1.		
DR	PRINTS; PR00507; N12N6MTFRASE.		
DR	PROSITE; PS00092; N6_MTASE; UNKNOWN 1.		
KW	Complete proteome; Hypothetical protein; Plasmid.		
SQ	SEQUENCE 1277 AA; 150972 MW; 0841DF4EDB5859F4 CRC64;		

Query Match 99.7%; Score 872; DB 2; Length 1277;

Best Local Similarity 99.4%; Pred. No. 1.2e-55;

Matches 173; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QITQHILASILESIIFILRFFIAYIENDIPFKILQENKLYRSISFYFFYDENTKK 60

Db 268 EITQHILASILESIIFILRFFIAYIENDIPFKILQENKLYRSISFYFFYDENTKK 327

QY 61 KLEYKKIITIFNLLDKSGDAIKFPVPFNGGLFSDKVKYLNNEGLLSISEIEILVKMLFF 120


```

Db      328 KLEYKKIITIFNLLDKGSDAIFKFPVNGGLFSEDKVKYLNNEGLLSISEEELIVKMLFF 387
QY      121 EEKNIKDEKFKVYSLRDPKSGFGLYETLLLEYDLRIADTTVHRIIEDGVYLIRTE 174
Db      388 EEKNIKDEKFKVYSLRDPKSGFGLYETLLLEYDLRIADTTVHRIIEDGVYLIRTE 441

RESULT 2
OS0667 BORBU
ID OS0667 BORBU PRELIMINARY; PRT; 1278 AA.
AC OS0667;
DT 01-JUN-1998 (T-EMBLrel. 06, Created)
DT 01-JUN-1998 (T-EMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Hypothetical protein BBH09.
GN OrderedLocuNames=BBH09;
OC Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp28-3.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=139;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hackey E.K.,
RA Winn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson D.L., Petersen J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A.,
RA Artiach P., Bowman K.C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi."
RL Nature 390:580-586 (1997).
DR EMBL; AE000784; AAC66000.1; -; Genomic_DNA.
DR PIR; B70236; B70236.
DR TIGR; BBH09; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR InterPro; IPR007409; DUF450.
DR InterPro; IPR002296; N12N6_mtfase.
DR Pfam; PF04313; HSDR N; 1.
DR PROSITE; PS00507; N12N6MTFRASE.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN 1.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 1278 AA; 150871 MW; 8A0B22DF166565C0 CRC64;

Query Match 95.3%; Score 834; DB 2; Length 1278;
Best Local Similarity 92.5%; Pred. No. 7.4e-53;
Matches 161; Conservative 11; Mismatches 2; Indels 0; Gaps 0;

QY 1 QITQHILASILEESIIFILRIFFIAYIEDNDIFKKILOENKLYRSSISFRYFFYDENTKK 60
Db 274 EITRHILASILEESIIFILRIFFIAYIEDNDIFKKILEENKLYRSSVSFRYFFYDENTKK 333
QY 61 KLEYKKIITIFNLLDKGSDAIFKFPVNGGLFSEDKVKYLNNEGLLSISEEELIVKMLFF 120
Db 334 KLGKXKIITIFNLLDKGSDAIFKFPVNGGLFAEDKVKYLNNEGLLSISEEELIVKILFF 393
QY 121 EEKNIKDEKFKVYSLRDPKSGFGLYETLLLEYDLRIADTTVHRIIEDGVYLIRTE 174
Db 394 EEKNIKDEKFKVYSLRDPKSGFGLYETLLLEYDLRIADTTVHRIEDGVYLIRTE 447

RESULT 3
Q5XYT2 BORGA
ID Q5XYT2 BORGA PRELIMINARY; PRT; 1194 AA.
AC Q5XYT2;

Query Match 93.8%; Score 821; DB 2; Length 1194;
Best Local Similarity 92.0%; Pred. No. 6.3e-52;
Matches 160; Conservative 10; Mismatches 4; Indels 0; Gaps 0;

QY 1 QITQHILASILEESIIFILRIFFIAYIEDNDIFKKILOENKLYRSSISFRYFFYDENTKK 60
Db 268 EITQILASILEESIIFILRIFFIAYIEDNDIFKKILEENKLYRSSISFRYFFYDENTKK 327
QY 61 KLEYKKIITIFNLLDKGSDAIFKFPVNGGLFSEDKVKYLNNEGLLSISEEELIVKMLFF 120
Db 328 KLGKXKIITIFNLLDKGSDAIFKFPVNGGLFAEDKVKYLNNEGLLSISEEELIVKILFF 387
QY 121 EEKNIKDEKFKVYSLRDPKSGFGLYETLLLEYDLRIADTTVHRIIEDGVYLIRTE 174
Db 388 EEKNIKDEKFKVYSLRDPKSGFGLYETLLLEYDLRIADTTVHRIEDGVYLIRTE 441

RESULT 4
Q5XY16 BORGA
ID Q5XY16 BORGA PRELIMINARY; PRT; 149 AA.
AC Q5XY16;
DT 25-OCT-2004 (T-EMBLrel. 28, Created)
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=BGP295;
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=290434;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBI;
RA Glockner G., Lehmann R., Romualdi A., Pradella S.,
RA Schulte-Spechtel U., Schilhabel M., Wilske B., Suhnel J., Platzner M.;
RA "Comparative analysis of the Borrelia garinii genome."
RT "Comparative analysis of the Borrelia garinii genome."
RL Nucleic Acids Res. 32:6038-6046(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.

```



```
RC STRAIN=PBI;
RA Gloeckner G., Schilhabel M., Lehmann R., Platzter M.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY722940; AAU86146.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 149 AA; 17344 MW; CB660C79D86F038E CRC64;

Query Match      53.7%; Score 470; DB 2; Length 149;
Best Local Similarity 86.3%; Pred. No. 5.1e-27;
Matches 88; Conservative 12; Mismatches 2; Indels 0; Gaps 0;

QY 73 LLDKGSDAIKFPVNGSLFSDKVKYLNNEGLSISEIEILVKMLFFFEKNIKDKFKVK 132
DB 1 MLDDKGSDAIKFISFNGSLFADKVKYLNKSLLSISEIEELVKLIIFPEKNIKDKFKVE 60

QY 133 YSRLDPKSPGELYETLLEYDLRIADTTVHRIIDGVYLIRTE 174
DB 61 YSKLDPKSPGELYETLLEYDLRIADTTVHRIVDGVYLIRTE 102

RESULT 5
Q5XYG5_BORGA
ID Q5XYG5_BORGA PRELIMINARY; PRT; 806 AA.
AC Q5XYG5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DR EMBL; AY722945; AAU86167.1; -; Genomic_DNA.
DE Hypothetical protein (Fragment).
GN ORFNames=BGP316;
OS Borrelia garinii PBI.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=290434;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBI;
RA Gloeckner G., Lehmann R., Romualdi A., Pradella S.,
RA Schulte-Spechtel U., Schilhabel M., Wilske B., Sühnel J., Platzter M.;
RT "Comparative analysis of the Borrelia garinii genome.";
RN Nucleic Acids Res. 32:6038-6046(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBI;
RA Gloeckner G., Schilhabel M., Lehmann R., Platzter M.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY722945; AAU86167.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR InterPro; IPR002296; N12N6_mtfase.
DR PRINTS; PR00507; N12N6MTFRASE.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
KW Hypothetical protein.
FT NON TER 806 806
SQ SEQUENCE 806 AA; 94261 MW; 2590359906792188 CRC64;

Query Match      33.1%; Score 290; DB 2; Length 806;
Best Local Similarity 77.5%; Pred. No. 5.2e-13;
Matches 55; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 104 LLSISEIEILVKMLFFFEKNIKDKFKVKYSRLDPKSPGELYETLLEYDLRIADTTVHRI 163
DB 1 MLSISEIEILIKLIFLEEKNIKDKFKVFKYSLDKPKIFGELYENLKYDLRIADTTTHRI 60

QY 164 IEDGVYLIRTE 174
DB 61 IEKGTTCIHT 71

RESULT 6
Q5HV91_CAMJR
ID Q5HV91_CAMJR PRELIMINARY; PRT; 1250 AA.
```

```
Q5HV91;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=CJ080789;
OS Campylobacter jejuni (strain RM1221).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=195099;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15660156; DOI=10.1371/journal.pbio.0030015;
RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
RA Ravel J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
RA Dodson R.J., Durkin A.S., Madupu R., Sullivan S.A., Shetty J.U.,
RA Ayodemi M.A., Shvartsbeyn A., Schatz M.C., Badger J.H., Fraser C.M.,
RA Nelson K.E.;
RT "Major structural differences and novel potential virulence mechanisms
from the genomes of multiple Campylobacter species.";
RL PLOS Biol. 3:72-85(2005).
DR EMBL; CP000025; AAW34576.1; -; Genomic_DNA.
DR TIGR; CJ080789; -;
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR InterPro; IPR007409; DUP450.
DR InterPro; IPR002296; N12N6_mtfase.
DR InterPro; IPR002052; N6_Mtase.
DR Pfam; PF04313; HSDR_N; 1.
DR PRINTS; PR00507; N12N6MTFRASE.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 1250 AA; 147549 MW; A367A4D74887B235 CRC64;

Query Match      26.5%; Score 232; DB 2; Length 1250;
Best Local Similarity 33.7%; Pred. No. 1.5e-08;
Matches 55; Conservative 34; Mismatches 68; Indels 6; Gaps 3;

QY 3 TCHLASILEESIIFILRIFFIATIEDNDIFKKLOENKLYRSSISFRYFF--YDENTKK 60
DB 260 TRADLKLIYENSLSYFIFLLFIATYPEDK--FEILEKHVKFKSKISLRLLENLQDESS 317

QY 61 KLEYKKIITFNLLDKGSDAIKFPVNGSLFSDKVKYLNNEGLSISEIEILVKMLFF 120
DB 318 SGCGELENIFNIYKNGKGFDPMPVNGSLFDESKTALLSTPKIFNDKOLKFLINQLNLF 377

QY 121 EEKNIKDEKFKVYSRLDPKSPGELYETLLEYDLRIADTTVHRI 163
DB 378 KOKNLSFKR--DYKTLVSVEHLGTIYEGLLSYFPEIANEDIYV 418

RESULT 7
Q9PPL7_CAMJE
ID Q9PPL7_CAMJE PRELIMINARY; PRT; 1250 AA.
AC Q9PPL7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Possible restriction/modification enzyme.
GN OrderedLocusNames=Cj0690c;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204; DOI=10.1038/35001088;
RA Parkhill J., Wren B.W., Mungall K.L., Ketley J.M., Churcher C.M.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagsels K., Karlyshev A.V., Moule S., Fallan M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
```

[illegible]

RESULT 9	
OS2000 HALSA	
ID OS2000 HALSA PRELIMINARY;	PRT; 706 AA.
AC OS2000;	
DT 01-JUN-1998 (TrEMBLrel. 06, Created)	
DT 01-JUN-1998 (TrEMBLrel. 06, last sequence update)	
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)	
DE Hypothetical protein.	
GN OrderedLocustNames=VNG5138C;	
OS Halobacterium salinarium (Halobacterium halobium) .	
OG Plasmid pNRC100.	
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;	
OC Halobacteriaceae; Halobacterium.	
OX NCBI TaxID=2242;	

[illegible]

```

Db 274 IHDSLLIYVRLIFVLYVABEAG--RELLDTSNIEYQSYSLSLKSQEIABELDSGDPKVR 331
QY 53 FYDENTKKLEKKIITIFNLLDKGS-----DAIKFPVNGGLFSED-----KVKY 98
Db 332 DWQDNQSLRD-----ELFTLNKSGSRGIPEDLYIPAYNGGLFRTDDEDDSAEAKF 386
QY 99 LNNE--GLGISIEIEILVQMLFFEEKHIDEK-FVKYSRLDPKSPGELYETLLEYDLRI 155
Db 387 LANHDVGDAYLAKVWELLTR-----SKNDGGGKIFVDYSSLDVRLHLSGYEGLLEYQLNV 442
QY 156 ADTTVRIIIEGCVLIRTE 174
Db 443 ADEPL--ALDDGYVSADE 459

RESULT 10
Q50N37_ENTHI PRELIMINARY; PRT; 1193 AA.
AC Q50N37;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=465.t00009;
OS Entamoeba histolytica HM-1:IMSS.
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=294381;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HM-1:IMSS;
RX PubMed=15729342; DOI=10.1038/nature03291;
RA Loftus B., Anderson I., Davies R., Alsmark U.C., Samuelson J.,
RA Amedeo P., Roncaglia P., Berriman M., Hirt R.P., Mann B.J., Nozaki T.,
RA Suh B., Pop M., Duchene M., Ackers J., Tannich E., Leippe M.,
RA Hofer M., Bruchhaus I., Willhoeft U., Bhattacharya A.,
RA Chillingworth T., Churcher C., Hance Z., Harris B., Harris D.,
RA Jagels K., Moule S., Mungall K., Ormond D., Squares R., Whitehead S.,
RA Quail M.A., Rabinowitsch E., Norbertczak H., Price C., Wang Z.,
RA Guillen N., Gilchrist C., Stroup S.E., Bhattacharya S., Lohia A.,
RA Foster P.G., Sichert-Ponten T., Weber C., Singh U., Mukherjee C.,
RA El-Sayed N.M., Petri W.A., Clark C.G., Embley T.M., Barrell B.,
RA Fraser C.M., Hall N.;
RT "The genome of the protist parasite Entamoeba histolytica."
RL Nature 433:865-868(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAFB01000388; EAL47296.1; -; Genomic_DNA.
KW Hydrolyase.
SQ SEQUENCE 2181 AA; 253911 MW; 4968693964C71AAE CRC64;

Query Match 12.5%; Score 109.5; DB 2; Length 2181;
Best Local Similarity 27.9%; Pred. No. 28;
Matches 50; Conservative 30; Mismatches 64; Indels 35; Gaps 11;

QY 28 EDNDIFPKLQENKLYRSSISFRYF---FYDENTKKLE--YKKII-----TIFNLLDKG 77
Db 58 EMNVIVKKISEENVFYMYRSEYETFLKFTTTNMKLSDOFYKEFVNKNIKVFNKIKE 117
QY 78 SD-----AIKFPVNGGLFSEDVKVYLNNEGLLSISEI---EEI-----LVKMLFFEE 122
Db 118 TWYTELILKFTTFYLNLF--SRIELVDEGKLFGLDFNVEEFKSKTVVNIKAKYIE 176
QY 123 KNKDEKFKVYSLRDPKSPGELYETLLEYDLRIAD---TTVHRII-EDGVYLI---RTE 174
Db 177 VPLVVEFIIIEINKID--GFNKIYEMIRLYDLKLLNSMLTSVSYFSNDNMVLLIEDQORTE 233

RESULT 11
Q510L7_ENTHI PRELIMINARY; PRT; 2181 AA.
AC Q510L7;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)

```

```

DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE Ubiquitin carboxyl-terminal hydrolase, putative.
GN ORFNames=109.t00017;
OS Entamoeba histolytica HM-1:IMSS.
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=294381;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HM-1:IMSS;
RX PubMed=15729342; DOI=10.1038/nature03291;
RA Loftus B., Anderson I., Davies R., Alsmark U.C., Samuelson J.,
RA Amedeo P., Roncaglia P., Berriman M., Hirt R.P., Mann B.J., Nozaki T.,
RA Suh B., Pop M., Duchene M., Ackers J., Tannich E., Leippe M.,
RA Hofer M., Bruchhaus I., Willhoeft U., Bhattacharya A.,
RA Chillingworth T., Churcher C., Hance Z., Harris B., Harris D.,
RA Jagels K., Moule S., Mungall K., Ormond D., Squares R., Whitehead S.,
RA Quail M.A., Rabinowitsch E., Norbertczak H., Price C., Wang Z.,
RA Guillen N., Gilchrist C., Stroup S.E., Bhattacharya S., Lohia A.,
RA Foster P.G., Sichert-Ponten T., Weber C., Singh U., Mukherjee C.,
RA El-Sayed N.M., Petri W.A., Clark C.G., Embley T.M., Barrell B.,
RA Fraser C.M., Hall N.;
RT "The genome of the protist parasite Entamoeba histolytica."
RL Nature 433:865-868(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAFB01000388; EAL47296.1; -; Genomic_DNA.
KW Hydrolyase.
SQ SEQUENCE 2181 AA; 253911 MW; 4968693964C71AAE CRC64;

Query Match 12.5%; Score 109.5; DB 2; Length 2181;
Best Local Similarity 27.9%; Pred. No. 28;
Matches 50; Conservative 30; Mismatches 64; Indels 35; Gaps 11;

QY 28 EDNDIFPKLQENKLYRSSISFRYF---FYDENTKKLE--YKKII-----TIFNLLDKG 77
Db 58 EMNVIVKKISEENVFYMYRSEYETFLKFTTTNMKLSDOFYKEFVNKNIKVFNKIKE 117
QY 78 SD-----AIKFPVNGGLFSEDVKVYLNNEGLLSISEI---EEI-----LVKMLFFEE 122
Db 118 TWYTELILKFTTFYLNLF--SRIELVDEGKLFGLDFNVEEFKSKTVVNIKAKYIE 176
QY 123 KNKDEKFKVYSLRDPKSPGELYETLLEYDLRIAD---TTVHRII-EDGVYLI---RTE 174
Db 177 VPLVVEFIIIEINKID--GFNKIYEMIRLYDLKLLNSMLTSVSYFSNDNMVLLIEDQORTE 233

RESULT 12
Q4XXL0_PLACH PRELIMINARY; PRT; 2337 AA.
AC Q4XXL0;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE Hypothetical protein (Fragment).
GN ORFNames=PC001007.02.0;
OS Plasmodium chabaudi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5825;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hall N., Karas M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C.,
RA Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT "A comprehensive survey of the Plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses."
RL Science 307:82-86(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

```

```

CC preliminary data.
DR EMBL; CAJ01002429; CAH78351.1; -; Genomic_DNA.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 2337 2337 1
SQ SEQUENCE 2337 AA; 289397 MW; 0DCA834991D6A8CA CRC64;

Query Match 12.4%; Score 108.5; DB 2; Length 2337;
Best Local Similarity 27.2%; Pred. No. 36;
Matches 41; Conservative 26; Mismatches 55; Indels 29; Gaps 6;

QY 5 HILASILERSIFILRIFFIAYIEDNDIFKKILOENKLYRSSI-SPRYFFYDENTKKLE 63
DB 1948 HNWVSIVEKLIYVNF-----NEI-KKILKKNIFKOVIRNFYFNFKLKXKDMQ 1998

QY 64 YKKIITIFNLIDKGSDAIKFPVFNGLFSEDKVKYL-----NNEGLLSI 107
DB 1999 KKYFQIYVNLVKKNSIINKLQIYVSLQYKRMKYLFFYNRKNTRKKLQNYTHLS- 2057

QY 108 SEIEEILV--KMLFFFEKNIKDKFKVYSRL 136
DB 2058 KKKHSIFTKWKAIFIKNNINDEKLAFNKL 2088

RESULT 13
Q54H63_DICDI
ID Q54H63_DICDI PRELIMINARY; PRT; 868 AA.
AC Q54H63;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DD80188518;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachobab J.A., Gloeckner G., Rajandream M.-A.,
RA Sugang R., Berriman M., Song J., Olsen R., Safranski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Louissege H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Plattner M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,
RA "The genome of the social amoeba Dictyostelium discoideum.";
RT Nature 0:0-0(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAFI01000219; EAL62554.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 868 AA; 105066 MW; D7DCD2A3EE249CC1 CRC64;

Query Match 12.3%; Score 107.5; DB 2; Length 868;
Best Local Similarity 24.7%; Pred. No. 15;
Matches 40; Conservative 36; Mismatches 55; Indels 31; Gaps 7;

QY 19 LRIFFIAYIE--DNDIF-----KKILOENKLYRSSIS--PRYFFYDENTKKLEYK--- 65
DB 284 LKEFFKSYIKSINNNSFKFENKNDIITNVLVKKYIKKLVKLLVYDENGKKKDKFKIPD 343

```

```

QY 66 -KII-----TIFNLIDKGSDAIKFPVFNGLFSEDKVKYLNNEGLLSISEIE 112
DB 344 LKEIESMKFTNQFNSTISDLITDGDDE---NYEKDI--QENLRILYSLFFIYQYDH 398

QY 113 ILVKMLFFFEKNIKDKFKVYSRLDPKSGFELYETLLLEYDLR 154
DB 399 LIGNLIFYKVKILNEKDFMKYNHLYPKLCPKMGIVGIYDYK 440

RESULT 14
Q5FM50_LACAC
ID Q5FM50_LACAC PRELIMINARY; PRT; 939 AA.
AC Q5FM50;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Methylase.
GN OrderedLocusNames=LBA0332;
OS Lactobacillus acidophilus.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1579;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NCFM;
RA Altermann E., Russell W.M., Azcarate-Perill M.A., Barrangou R.,
RA Luck B.L., McAuliffe O., Southern N., Dobson A., Duong T., Callanan M.,
RA Lick S., Hamrick A., Cano R., Klammer T.R.;
RT "Complete genome sequence of the probiotic lactic acid bacterium
RT Lactobacillus acidophilus NCFM.";
RL Proc. Natl. Acad. Sci. U.S.A. 102:3906-3912(2005).
DR EMBL; CP000033; AAV42224.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR InterPro; IPR002296; N12N6_mtfase.
DR PRINTS; PR00507; N12N6MTRASE.
DR Complete proteome; Methyltransferase.
KW SEQUENCE 939 AA; 108344 MW; 4635DB98E12A08C CRC64;

Query Match 12.3%; Score 107.5; DB 2; Length 939;
Best Local Similarity 25.0%; Pred. No. 17;
Matches 42; Conservative 25; Mismatches 52; Indels 49; Gaps 8;

QY 19 LRIFFIAYIEDNDIFKKILOENKLYRSSISFRYFFYD--ENTKKLEYKKIIFNLID- 75
DB 208 VRIVFCLYAEADAGIFGK-----KNMFHDYLEEFDAHRMRKALINLFKVLDT 253

QY 76 --KGSD-----AIKFPVFNGLFSEDKVKYLNNEGLLSISEIEILVKMLFPEKN 124
DB 254 KVXDRDPYLEDDSPKLAQFPYVNGMFSDE-----NIEIPPTDELKDLLL-----SK 301

QY 125 IKDEKFKVYSRLDPKSGFELYETLLLEYDLR-----IADTVHRIIE 165
DB 302 ASDS--FDWSEISFTIFGAVFESTLAPDTRRQGMHYTSIENHKVID 347

RESULT 15
Q8FSZ4_COREF
ID Q8FSZ4_COREF PRELIMINARY; PRT; 1524 AA.
AC Q8FSZ4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=CE2898;
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]

```

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Y8-314 / AJ 12310 / DSM 44549 / JCM 11189;
RX MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of *Corynebacterium*
RT efficiens";
RL Genome Res. 13:1572-1579 (2003).
DR EMBL; BA000035; BAC19708.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 1524 AA; 171619 MW; 6FCE5647350F256E CRC64;

Query Match 12.2%; Score 106.5; DB 2; Length 1524;
Best Local Similarity 20.8%; Pred. No. 33;
Matches 42; Conservative 37; Mismatches 58; Indels 65; Gaps 8;

QY 10 ILRESIIILRIFFIAYIE---DNDIFKKILQENKLYRSSISPRYPYDE----- 56
Db 281 LAKQSKLYLRIILFLFAEASPELEILFTGTPE-----YDEGYGLSLREL 326

QY 57 -----NTKKLE-YKKIITIFNLIDKGD-----AIKFPVFNGLF 91
Db 327 ILTPPVTONARSGTHLYDSQLFLHLVDQNDPRDSNPDPDPAGEGLQFRNLSADLF 386

QY 92 SE-----DKVYNNEGLLSISEIEILVKMLFFBEKNIKDEKFKYKSLDPKSPGELY 145
Db 387 QPKATNIIDTVK-LSNEAL-----NTVLENLLLSKEQSGKRGFISYATLGVTELQVY 439

QY 146 ETLLDYDLRIADTVVHRIIEDG 167
Db 440 EGLMSYTGFIAREDLFEVAKNG 461

Search completed: January 24, 2006, 19:56:50
Job time : 39.3889 secs

This Page Blank (uspto)


```

US-08-015-986A-14
; Sequence 14, Application US/08015986A
; Patent No. 5532123
; GENERAL INFORMATION:
; APPLICANT: Schlössinger, Joseph
; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASE-GAMMA
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/015,986A
; FILING DATE: 10-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 373 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-015-986A-14

Query Match 11.1%; Score 97; DB 1; Length 373;
Best Local Similarity 21.8%; Pred. No. 0.063;
Matches 48; Conservative 36; Mismatches 50; Indels 86; Gaps 11;

Qy 26 YIEDNDIF----KKILOENKLYRSISFRYPFFYDENTKKKLEYKKII-----68
Db 123 YCFDADRFSFEAEVKGKGLRALSILF-----EVGTEENLDFKAIIDGVESVSRFGKQA 177
Qy 69 -----TIFNLIDKGSDAIKFPVNGGLFS---EDKVYLNNEGLLSISEIE-----EIL- 114
Db 178 ALDPFILLNLPNSTD--KYIYNGSLTSPCTDTVDWIVFKDTVSISQSLAVFCEVLT 235
Qy 115 -----VKMLFFEEKNIKDEKFKVYSR-----LDPKSFGE 143
Db 236 MQQSGYVLMYLDYLNQNFREQY-KFSRQVFSYTGKEIHEAVCSSEPNVQADPENYTS 294
Qy 144 L-----YETLLE-----YDLRIADTTVHRIEDG 167
Db 295 LLVTWPRVYVDTMIKFAVLVYQQLDGEDQTKHEFLTDG 334

RESULT 3
US-08-446-363-14
; Sequence 14, Application US/08446363
; Patent No. 5891700
; GENERAL INFORMATION:
; APPLICANT: Schlössinger, Joseph
; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASE-GAMMA
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/015,986A
; FILING DATE: 10-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 373 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-015-986A-14

Query Match 11.1%; Score 97; DB 1; Length 373;
Best Local Similarity 21.8%; Pred. No. 0.063;
Matches 48; Conservative 36; Mismatches 50; Indels 86; Gaps 11;

Qy 26 YIEDNDIF----KKILOENKLYRSISFRYPFFYDENTKKKLEYKKII-----68
Db 123 YCFDADRFSFEAEVKGKGLRALSILF-----EVGTEENLDFKAIIDGVESVSRFGKQA 177
Qy 69 -----TIFNLIDKGSDAIKFPVNGGLFS---EDKVYLNNEGLLSISEIE-----EIL- 114
Db 178 ALDPFILLNLPNSTD--KYIYNGSLTSPCTDTVDWIVFKDTVSISQSLAVFCEVLT 235
Qy 115 -----VKMLFFEEKNIKDEKFKVYSR-----LDPKSFGE 143
Db 236 MQQSGYVLMYLDYLNQNFREQY-KFSRQVFSYTGKEIHEAVCSSEPNVQADPENYTS 294
Qy 144 L-----YETLLE-----YDLRIADTTVHRIEDG 167
Db 295 LLVTWPRVYVDTMIKFAVLVYQQLDGEDQTKHEFLTDG 334

RESULT 4
US-08-015-973-1
; Sequence 1, Application US/08015973
; Patent No. 5604094
; GENERAL INFORMATION:
; APPLICANT: Schlössinger, Joseph
; TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASE-BETA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```


COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/015,973
FILING DATE: 10-FEB-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7683-021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2308 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-015-973-1

Query Match 11.1%; Score 97; DB 1; Length 2308;
Best Local Similarity 21.8%; Pred. No. 0.76;
Matches 48; Conservative 36; Mismatches 50; Indels 86; Gaps 11;
QY 26 YIEDNDIP-----KKILOENKLYRSSISFRYFFDENTKKKLEYKKII-----68
DB 156 YCFDADRFSFEAVKGGKLRALSILF-----EVGTEENLDFKAIIDGVESVSRFGKQA 210
QY 69 -----TIFNLLDKGSDAIKPPVNGGLFS---EDKVKYLNNEGLLSISIE-----EIL- 114
DB 211 ALDPFILLNLLPNSTD--KYIYNGSLTSPCTDTVDWIVFKDVTVSISQSALAVFCEVLT 268
QY 115 -----VKMLFFEEKNIKDEKFKVYSR-----LDPKSFG 143
DB 269 MOOSGYVMDLYQNNFREQQY-KFSQVFSSTGKEEIHAEVCSPEPVQADPENYTS 327
QY 144 L-----YETLLE-----YDLRIADTTVHRIEDG 167
DB 328 LLVTWERPRVYDTMIKFAVLVQQLDGEDQTKHEFLTGD 367

RESULT 5
US-08-448-164-1
Sequence 1, Application US/08448164
Patent No. 5925536
GENERAL INFORMATION:
APPLICANT: Schlensing, Joseph
TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
TITLE OF INVENTION: PHOSPHATASE-BETA
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,164
FILING DATE: 24-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/015,973
FILING DATE: 10-FEB-1993
ATTORNEY/AGENT INFORMATION:

NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7683-021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2308 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-448-164-1
Query Match 11.1%; Score 97; DB 1; Length 2308;
Best Local Similarity 21.8%; Pred. No. 0.76;
Matches 48; Conservative 36; Mismatches 50; Indels 86; Gaps 11;
QY 26 YIEDNDIP-----KKILOENKLYRSSISFRYFFDENTKKKLEYKKII-----68
DB 156 YCFDADRFSFEAVKGGKLRALSILF-----EVGTEENLDFKAIIDGVESVSRFGKQA 210
QY 69 -----TIFNLLDKGSDAIKPPVNGGLFS---EDKVKYLNNEGLLSISIE-----EIL- 114
DB 211 ALDPFILLNLLPNSTD--KYIYNGSLTSPCTDTVDWIVFKDVTVSISQSALAVFCEVLT 268
QY 115 -----VKMLFFEEKNIKDEKFKVYSR-----LDPKSFG 143
DB 269 MOOSGYVMDLYQNNFREQQY-KFSQVFSSTGKEEIHAEVCSPEPVQADPENYTS 327
QY 144 L-----YETLLE-----YDLRIADTTVHRIEDG 167
DB 328 LLVTWERPRVYDTMIKFAVLVQQLDGEDQTKHEFLTGD 367

RESULT 6
US-08-081-929-2
Sequence 2, Application US/08081929
Patent No. 6160090
GENERAL INFORMATION:
APPLICANT: Schlensing, Joseph
APPLICANT: Barnea, Gilad
APPLICANT: Grumet, Martin H.
APPLICANT: Margolis, Richard U.
TITLE OF INVENTION: A NEW CLASS OF RPTASES: THEIR
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/081,929
FILING DATE: 23-JUN-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30742
REFERENCE/DOCKET NUMBER: 7683-041-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2308 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-081-929-2

Query Match      11.1%; Score 97; DB 2; Length 2308;
Best Local Similarity 21.8%; Pred. No. 0.76;
Matches 48; Conservative 36; Mismatches 50; Indels 86; Gaps 11;

QY 26 YIEDNDIF----KKILOENKLYRSSISFRYFFYDENTKKKLEYKKII-----68
Db 156 YCPDADRFSFEAEVKGKLRALSILF-----EVGTEENLDKFAIIDGVESVSRFGKQA 210
QY 69 -----TIFNLLDKGSDAIKFPVNGGLFS---EDVKYLNNEGLLSTSEIE-----EIL- 114
Db 211 ALDPFILLNLPNSTD--KYIYNGSLTSPPTDVTDMIVFKDVTVSISESQLAVFCEVLT 268
QY 115 -----VKMLFFEEKNIKDEKFKVYSR-----LDPKSFGE 143
Db 269 MQQSGYVLMLDYLNQNFREQY-KFSRQVFSSTGKEEIHAEVCSSEPENVQADPENYTS 327
QY 144 L-----YETLLE-----YDLRIADTTVHRIIEDG 167
Db 328 LLVTWERPRVYDTMIEKFAVLVYQQLDGEDQTKHEFLTGD 367

RESULT 7
US-10-000-954-2
; Sequence 2, Application US/10000954
; Patent No. 6803452
; GENERAL INFORMATION:
; APPLICANT: Schlusseringer, Joseph
; Barnea, Gilad
; Grumet, Martin H.
; Margolis, Richard U.
; TITLE OF INVENTION: A NEW CLASS OF RPTASES: THEIR
; STRUCTURAL DOMAINS AND LIGANDS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/000,954
; FILING DATE: 04-Dec-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/644,293
; FILING DATE: 23-Aug-2000
; APPLICATION NUMBER: 08/081,929
; FILING DATE: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30742
; REFERENCE/DOCKET NUMBER: 7683-041-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2308 amino acids
```

```
;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-000-954-2

Query Match      11.1%; Score 97; DB 2; Length 2308;
Best Local Similarity 21.8%; Pred. No. 0.76;
Matches 48; Conservative 36; Mismatches 50; Indels 86; Gaps 11;

QY 26 YIEDNDIF----KKILOENKLYRSSISFRYFFYDENTKKKLEYKKII-----68
Db 156 YCPDADRFSFEAEVKGKLRALSILF-----EVGTEENLDKFAIIDGVESVSRFGKQA 210
QY 69 -----TIFNLLDKGSDAIKFPVNGGLFS---EDVKYLNNEGLLSTSEIE-----EIL- 114
Db 211 ALDPFILLNLPNSTD--KYIYNGSLTSPPTDVTDMIVFKDVTVSISESQLAVFCEVLT 268
QY 115 -----VKMLFFEEKNIKDEKFKVYSR-----LDPKSFGE 143
Db 269 MQQSGYVLMLDYLNQNFREQY-KFSRQVFSSTGKEEIHAEVCSSEPENVQADPENYTS 327
QY 144 L-----YETLLE-----YDLRIADTTVHRIIEDG 167
Db 328 LLVTWERPRVYDTMIEKFAVLVYQQLDGEDQTKHEFLTGD 367

RESULT 8
US-09-816-703A-2
; Sequence 2, Application US/09816703A
; Patent No. 6455026
; GENERAL INFORMATION:
; APPLICANT: Melcher, Thorsten
; APPLICANT: Mueller, Sabine
; APPLICANT: Chin, Daniel
; TITLE OF INVENTION: Use of Protein Tyrosine Phosphatase Zeta as a Biomolecular Target
; TITLE OF INVENTION: Treatment and Visualization of Brain Tumors
; FILE REFERENCE: 262/235 AGY
; CURRENT APPLICATION NUMBER: US/09/816,703A
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 2314
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Reference
; LOCATION: (1)..(2314)
; OTHER INFORMATION: Krueger, N.X. and Saito, H.: A human transmembrane protein-tyrosi
; OTHER INFORMATION: ne-phosphatase, PTP zeta, is expressed in brain and has an N-term
; OTHER INFORMATION: inal receptor domain homologous to carbonic anhydrases; Proc. Nat
; OTHER INFORMATION: l. Acad. Sci. USA 89 (16), 7417-7421 (1992)
; NAME/KEY: Reference
; LOCATION: (1)..(2314)
; OTHER INFORMATION: Levy, J.B., et al.: The cloning of a receptor-type protein tyrosi
; OTHER INFORMATION: ne phosphatase expressed in the central nervous system; J. Biol.
; OTHER INFORMATION: Chem. 268 (14), 10573-10581 (1993)
; NAME/KEY: SIGNAL
; LOCATION: (1)..(24)
; OTHER INFORMATION: By similarity
; NAME/KEY: DOMAIN
; LOCATION: (25)..(1635)
; OTHER INFORMATION: Extracellular (potential)
; NAME/KEY: misc_feature
; LOCATION: (25)..(2314)
; OTHER INFORMATION: Mature chain; protein-tyrosine phosphatase zeta
; NAME/KEY: DOMAIN
; LOCATION: (34)..(302)
; OTHER INFORMATION: Carbonic-anhydrase like
; NAME/KEY: SITE
; LOCATION: (105)..(105)
```

```
; OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)
; NAME/KEY: SITE
; LOCATION: (134)..(134)
; OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)
; NAME/KEY: SITE
; LOCATION: (223)..(223)
; OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)
; NAME/KEY: SITE
; LOCATION: (232)..(232)
; OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)
; NAME/KEY: DOMAIN
; LOCATION: (312)..(406)
; OTHER INFORMATION: Fibronectin Type-III
; NAME/KEY: SITE
; LOCATION: (324)..(324)
; OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)
; NAME/KEY: SITE
; LOCATION: (381)..(381)
; OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)
; NAME/KEY: SITE
; LOCATION: (497)..(497)
; OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)
; NAME/KEY: SITE
; LOCATION: (501)..(501)
; OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)
; NAME/KEY: SITE
; LOCATION: (552)..(552)
; OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)
; NAME/KEY: BINDING
; LOCATION: (587)..(587)
; OTHER INFORMATION: Chondroitin sulfate (potential)
; NAME/KEY: SITE
; LOCATION: (602)..(602)
; OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)
; NAME/KEY: SITE
; LOCATION: (629)..(629)
; OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)
; NAME/KEY: BINDING
; LOCATION: (637)..(637)
; OTHER INFORMATION: Chondroitin sulfate (potential)
; NAME/KEY: SITE
; LOCATION: (677)..(677)
; OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)
; NAME/KEY: VARSPLIC
; LOCATION: (755)..(1614)
; OTHER INFORMATION: Splicing variant; missing (in short isoform)
; NAME/KEY: BINDING
; LOCATION: (997)..(997)
; OTHER INFORMATION: Chondroitin sulfate (potential)
; NAME/KEY: SITE
; LOCATION: (1017)..(1017)
; OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)
; NAME/KEY: SITE
; LOCATION: (1050)..(1050)
; OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)
; NAME/KEY: SITE
; LOCATION: (1082)..(1082)
; OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)
; NAME/KEY: SITE
; LOCATION: (1122)..(1122)
; OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)
; NAME/KEY: SITE
; LOCATION: (1456)..(1456)
; OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)
; NAME/KEY: BINDING
; LOCATION: (1548)..(1548)
; OTHER INFORMATION: Chondroitin sulfate (potential)
; NAME/KEY: BINDING
; LOCATION: (1550)..(1550)
; OTHER INFORMATION: Chondroitin sulfate (potential)
; NAME/KEY: SITE
; LOCATION: (1561)..(1561)
; OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)
```

```
; NAME/KEY: SITE
; LOCATION: (1617)..(1617)
; OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)
; NAME/KEY: TRANSMEM
; LOCATION: (1636)..(1661)
; OTHER INFORMATION: Transmembrane region; potential
; NAME/KEY: DOMAIN
; LOCATION: (1662)..(2314)
; OTHER INFORMATION: Cytoplasmic (potential)
; NAME/KEY: CONFLICT
; LOCATION: (1722)..(1728)
; OTHER INFORMATION: Missing (in ref. 2)
; NAME/KEY: DOMAIN
; LOCATION: (1744)..(1997)
; OTHER INFORMATION: Protein-tyrosine phosphatase
; NAME/KEY: ACT SITE
; LOCATION: (1932)..(1932)
; OTHER INFORMATION: Active site; by similarity
; NAME/KEY: DOMAIN
; LOCATION: (1998)..(2314)
; OTHER INFORMATION: Protein-tyrosine phosphatase
; NAME/KEY: misc feature
; LOCATION: (2222)..(2222)
; OTHER INFORMATION: Ancestral active site
; US-09-816-703A-2

Query Match 11.1%; Score 97; DB 2; Length 2314;
Best Local Similarity 21.8%; Pred. No. 0.7%;
Matches 48; Conservative 36; Mismatches 50; Indels 86; Gaps 11;

QY 26 YIEDNIF---KKIIQENKLYRSSISFRFFVDENTKKKLEYKKII-----68
Db 156 YCFDADFSSFEBAVKGKRLSILF-----EVGTEENLDPKAITDGVESVSRFGKQA 210
QY 69 -----TIFNLLDKGSDAIKFPVFGGLFS---BDKVYLNNEGLLSTSEIE---BIL- 114
Db 211 ALDPFILLNLLPNSTD--KYIYNGSLTSPCTDTVDWIVFKDTVSISESQLAVFCBVL 269
QY 115 -----VKMLFFPEEKIKDEKFKVYSR-----LDPKSFG 143
Db 269 MQQSGYVWMLDYLNQNNFREQY-KFSRQVSSYTGKEEIHAEVCSSEPNVQADPENYTS 327
QY 144 L-----YETLLE-----YDLRIADTTVHRIEDG 167
Db 328 LLVTWBRPRVYDTMIEKFAVLQYQQLDGEDQTKHEFLTDG 367

RESULT 9
US-09-134-001C-4848
; Sequence 4848, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynm Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCES: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4848
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-4848

Query Match 11.0%; Score 96; DB 2; Length 404;
Best Local Similarity 26.4%; Pred. No. 0.089;
Matches 43; Conservative 25; Mismatches 51; Indels 44; Gaps 9;
```

```
Qy 26 YIEDNDIFKK-----ILQENKLYRSSISFRYFFYDENTKKKLEYKKIITIFNLLDKGSDAI 81
Db 154 YIMTSDINHKTOTIEYFKQHKYFNDYDANHIFFKQDNIVALSEEGKLV-----LNRDGHIM 208
Qy 82 KFPVFNGLP-----SEDKVKY--LNNEGLLSISEIIEILVKML--PFEEKNI 125
Db 209 ETPNGNGGVFKSLKAGYLDKMQQDHVKYIPLNN-----IDNVLVKVLDPLFAGFTV 260
Qy 126 KDEKFVKYSRLDPK---SFGELY-----ETLLEY---DLRIAD 157
Db 261 TQSKDITSKTIQPKDSESVGRLVNVDCKDTVILEYSELDTDIAN 303

RESULT 10
US-09-710-279-3256
; Sequence 3256, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3256
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-3256

Query Match 10.9%; Score 95.5; DB 2; Length 331;
Best Local Similarity 26.0%; Pred. No. 0.077;
Matches 40; Conservative 24; Mismatches 49; Indels 41; Gaps 8;

Qy 26 YIEDNDIFKK-----ILQENKLYRSSISFRYFFYDENTKKKLEYKKIITIFNLLDKGSDAI 81
Db 145 YIMTSDINHKTOTIEYFKQHKYFNDYDANHIFFKQDNIVALSEEGKLV-----LNRDGHIM 199
Qy 82 KFPVFNGLP-----SEDKVKY--LNNEGLLSISEIIEILVKML--PFEEKNI 125
Db 200 ETPNGNGGVFKSLKAGYLDKMQQDHVKYIIFLNN-----IDNVLVKVLDPLFAGFTV 251
Qy 126 KDEKFVKYSRLDPK---SFGELY-----ETLLEY 151
Db 252 TQSKDITSKTIQPKDSESVGRLVNVDCKDTVLEY 285

RESULT 11
US-09-107-532A-5494
; Sequence 5494, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
```

```
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5494:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...585
SEQUENCE DESCRIPTION: SEQ ID NO: 5494:
US-09-107-532A-5494

Query Match 10.9%; Score 95.5; DB 2; Length 585;
Best Local Similarity 23.0%; Pred. No. 0.17;
Matches 38; Conservative 26; Mismatches 52; Indels 49; Gaps 8;

Qy 34 KKILQENKLYRSSISP-RYFFYDENTKKK-----LEYKKIITIFNLLDKG----S 78
Db 117 EKVLINLVHNSAFAIQKSIRYSENKLYKIVFLKKEAITDYGRKVAIVYELKKNIPGCD 176
Qy 79 DAIKFP--VFGGLPSEDKVKYLNNEGLLSISEIIEILVKMLPFEEKN-----I 125
Db 177 DNVNASNRMPFGG-YKSDAVIINEDNMLDITELP----IDFTWTCRNQCSDSGREYELV 232
Qy 126 KDEKFVKYSRLDPK-----SFGELYETLLEYDL 153
Db 233 NETDFVKLIKNDCKEEMKQFKOCLFDSDSMFNEIYERLLTIDM 277

RESULT 12
US-09-605-703B-604
; Sequence 604, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 604
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-605-703B-604
```

Query Match 10.7%; Score 93.5; DB 2; Length 622;
Best Local Similarity 22.5%; Pred. No. 0.3;
Matches 27; Conservative 22; Mismatches 54; Indels 17; Gaps 2;

QY 64 YKKIITIFNLDDK-----GSDAIKFPVFNGLPSEDKVKYLNNEGLLSI 107
DB 6 YDSQLLFTLVKGGHPTDAKAVAFDAEAGEGLHFRNLSADLFLPAATELIDRVG-LSN 64

QY 108 SEIBEILVKMLFPEEKNIKDEKFKYKSRLOPKSGELYETILLEYDLRIADTTVHRIIEDG 167
DB 65 EALNKNVLELLSRVQSGKDRGFISYATLGTVELGQVYEGLSMTGTGFIQBDLFEVAPHG 124

RESULT 13
US-09-605-703B-600
; Sequence 600, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oekar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; FILE OF INVENTION: PROTEINS
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 600
; LENGTH: 1218
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-605-703B-600

Query Match 10.7%; Score 93.5; DB 2; Length 1218;
Best Local Similarity 22.5%; Pred. No. 0.74;
Matches 27; Conservative 22; Mismatches 54; Indels 17; Gaps 2;

QY 64 YKKIITIFNLDDK-----GSDAIKFPVFNGLPSEDKVKYLNNEGLLSI 107
DB 27 YDSQLLFTLVKGGHPTDAKAVAFDAEAGEGLHFRNLSADLFLPAATELIDRVG-LSN 85

QY 108 SEIBEILVKMLFPEEKNIKDEKFKYKSRLOPKSGELYETILLEYDLRIADTTVHRIIEDG 167
DB 86 EALNKNVLELLSRVQSGKDRGFISYATLGTVELGQVYEGLSMTGTGFIQBDLFEVAPHG 145

RESULT 14
US-09-800-170-18
; Sequence 18, Application US/09800170
; Patent No. 6481667
; GENERAL INFORMATION:
; APPLICANT: Kinsella, Todd
; TITLE OF INVENTION: IN VIVO PRODUCTION OF CYCLIC PEPTIDES
; FILE REFERENCE: A-68614-1/DJB/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/800,170
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/187,130
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-09-800-170-18

Query Match 10.6%; Score 93; DB 2; Length 389;
Best Local Similarity 19.5%; Pred. No. 0.18;
Matches 43; Conservative 41; Mismatches 76; Indels 60; Gaps 7;

QY 5 HILASILEESIIIFILRIFFIAYIEDN-----DIPKK----- 35
DB 179 YIEVPLEDFLLLLRFGIVSRLEKSTLKLISGKRNLELFRKHVGFDTSEKAKALDELISK 238

QY 36 -----ILQENKLYRSSISPRYFFYDENTKKLKYKKIITIFNLDDKGS--AIKFPV 85
DB 239 AKESERYPILEE--LRLGLLFGFTRNELRIENPTVVLMEILERIERSGNLAEKIAV 296

QY 86 FNGCLPSEDKVKYLNNEGLLSISEIBEILVKMLFPEEKNIKDEKFKYKSRLOPKSGELY 145
DB 297 LEGRIKEENYLRILEEGLIENGKLTGKLELVWRNREFDSKDVYVR-----NIV 349

QY 146 ETI-----LEYDLRIADTT--VHRIEDGVVLI RT 173
DB 350 ENLVFLPVEKVERIEYGVYDVTTHFNFVANGILVHNT 389

RESULT 15
US-09-107-532A-6196
; Sequence 6196, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6196:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 408 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...408
; SEQUENCE DESCRIPTION: SEQ ID NO: 6196:
US-09-107-532A-6196

Query Match 10.6%; Score 93; DB 2; Length 408;

Best Local Similarity 22.6%; Pred. No. 0.19;
Matches 51; Conservative 30; Mismatches 47; Indels 98; Gaps 12;

Qy	10	ILBESIIFILRIPIFIAYIEDNDIFKKILOENKLYRSSISFRFYDENTKKLEYKKIIT	69
Db	14	VLLKSLHVCCLKPIWVFNSFSIIE-----YRSDRT-TYFFYD-----YKCLLT	57
Qy	70	IFNLLDKGSDAIK-----FPVFNGGLFSE-----DKVKY-----	98
Db	58	-----DNIKKTKDFHFTGD-FNEMFLLYMFKTRYVYLFGPFRANNIDKVFFKLKMN	107
Qy	99	-----LNNEGLLSISEIEIILVKMLFPEEKNIKD---EKFKVYS-----	134
Db	108	NLNVAMADRERLYNSLQNLTLYSLGDIRDILILVHYFFTGKIEDLFHEPLIEYTENLSKT	167
Qy	135	-----RLDPKSGELYET-LLEY-----DLRIADTTVHRI	163
Db	168	IEQIKIDNLLSQNYDPEIYLFYENKILEYVKNGDIRNLENMVFNL	213

Search completed: January 24, 2006, 20:02:43
Job time : 11.7267 secs

Qy	16	IFILRIFTAYTENDNDIFKKILOENLYRSIS-----FRYFFYDENTKKGLBYKKI	67
Db	188	LYLVRLLPCLPAEDTITTEKSLFQYETKTLEDGSLAHINTFLFVLNTPQKKLN-	246
Qy	68	ITIFNLDKGSDAIKPPVFNGLPSE-----DKVKYLANNEGLLSISEITEILVRMLFF	120
Db	247	-----LDEHLAA--FPYINGKLFEEPLPAPQOKAM---REALDLCSLD	286
Qy	121	EENKINDEKFKYKSLRDPKSGELYETILLE	150
Db	287	-----WSRISPAIFGSLFOSIMD	304

RESULT 6

```

US-10-437-963-187688
; Sequence 187688, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 187688
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(392)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_84367C.1.pgp
US-10-437-963-187688

```

Query Match	11.3%	Score 99;	DB 4;	Length 392;
Best Local Similarity	28.3%	Pred. No. 1.1;		
Matches	34;	Conservative 27;	Mismatches 41;	Indels 18; Gaps 5;
QY	10	ILEESIFILRIFFIAYIEDNDIPKQLOENKLYRSSISFRYPFYDENTKKKLEYKKI--	67	
Db	83	VLINYVLLSSLVMPMISFL---IPKGVLONDIYRS---RPFWKDDHKKKKYRLIRW	134	
QY	68	-TIFNLLDKGSDAI-KFPVFNGLPSEDKKVYLNNEGLLSISIEISILVMGLPFEKNI	125	
Db	135	DIICQXKIDGGFVHNLEQVONKLLSKILKFINEOG-----VWKDLKKYFYQSI	188	

RESULT 7

US-09-801-368-166	Sequence 166, Application US/09801368	?	PRIOR FILING DATE: 2000-03-21
?	?	?	PRIOR APPLICATION NUMBER: 60/206,848
?	Patent No. US20020128250A1	?	PRIOR FILING DATE: 2000-05-23
?	?	?	PRIOR APPLICATION NUMBER: 60/207,727
?	GENERAL INFORMATION:	?	PRIOR FILING DATE: 2000-05-26
?	APPLICANT: Busby, Robert	?	PRIOR APPLICATION NUMBER: 60/230,335
?	APPLICANT: Cali, Brian	?	PRIOR FILING DATE: 2000-09-06
?	APPLICANT: Hecht, Peter	?	PRIOR APPLICATION NUMBER: 60/230,347
?	APPLICANT: Holtzman, Doug	?	PRIOR FILING DATE: 2000-09-09
?	APPLICANT: Madden, Kevin	?	PRIOR APPLICATION NUMBER: 60/242,578
?	APPLICANT: Maxon, Mary	?	PRIOR FILING DATE: 2000-10-23
?	APPLICANT: Milne, Todd	?	PRIOR APPLICATION NUMBER: 60/253,625
?	APPLICANT: No. US20020128250A1man, Thea	?	PRIOR FILING DATE: 2000-11-27
?	APPLICANT: Royer, John	?	PRIOR APPLICATION NUMBER: 60/257,931
?	APPLICANT: Salama, Sofie	?	PRIOR FILING DATE: 2000-12-22

RESULTS

US-10-282-122A-53613
 ; Sequence 53613, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Karl
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELI/PRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22

; ORGANISM: Homo sapiens
US-10-723-860-951

Query Match 11.1%; Score 97; DB 5; Length 1005;
Best Local Similarity 21.8%; Pred. No. 5.3;
Matches 48; Conservative 36; Mismatches 50; Indels 86; Gaps 11;

QY 26 YIEDNDIF---KKILQENKLYRSSISFRFFYDENTKKLEYKKII-----68
Db 156 YCFDADRFSFEAEVKGKLRALSILF-----EVGTEENLDFKAIIDGVESVSFRGKQA 210

QY 69 -----TIFNLLDKGSDAIFPVFNGGLFS---EDKVYLNNEGLLSISETE-----EIL- 114
Db 211 ALDPFILLNLLPNSTD--KYIYNGSLTSPCTDVTVDWIVFKDVTVSISESQLAVFCEVLT 268

QY 115 -----VKMLFFPEEKNIKDEKFKVYSR-----LDPKSFGE 143
Db 269 MQQSGYVLMYLDLQNNFREQQY-KFSRQVFSYTGKEEIIHEAACRSEPNVQADPENYTS 327

QY 144 L-----YETLLE-----YDLRIADTTVHRIEDG 167
Db 328 LLVTWPRVVDYTMIEKFAVLVYQQLDGEDQTKHEFLTGD 367

RESULT 12

US-10-732-923-19145
; Sequence 19145, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 19145
; LENGTH: 1005
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-732-923-19145

Query Match 11.1%; Score 97; DB 5; Length 1005;
Best Local Similarity 21.8%; Pred. No. 5.3;
Matches 48; Conservative 36; Mismatches 50; Indels 86; Gaps 11;

QY 26 YIEDNDIF---KKILQENKLYRSSISFRFFYDENTKKLEYKKII-----68
Db 156 YCFDADRFSFEAEVKGKLRALSILF-----EVGTEENLDFKAIIDGVESVSFRGKQA 210

QY 69 -----TIFNLLDKGSDAIFPVFNGGLFS---EDKVYLNNEGLLSISETE-----EIL- 114
Db 211 ALDPFILLNLLPNSTD--KYIYNGSLTSPCTDVTVDWIVFKDVTVSISESQLAVFCEVLT 268

QY 115 -----VKMLFFPEEKNIKDEKFKVYSR-----LDPKSFGE 143
Db 269 MQQSGYVLMYLDLQNNFREQQY-KFSRQVFSYTGKEEIIHEAACRSEPNVQADPENYTS 327

QY 144 L-----YETLLE-----YDLRIADTTVHRIEDG 167
Db 328 LLVTWPRVVDYTMIEKFAVLVYQQLDGEDQTKHEFLTGD 367

RESULT 13

US-10-295-027-1264
; Sequence 1264, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.

; APPLICANT: Glynnne, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1264
; LENGTH: 1113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1264

Query Match 11.1%; Score 97; DB 4; Length 1113;
Best Local Similarity 21.8%; Pred. No. 6.1;
Matches 48; Conservative 36; Mismatches 50; Indels 86; Gaps 11;

QY 26 YIEDNDIF---KKILQENKLYRSSISFRFFYDENTKKLEYKKII-----68
Db 156 YCFDADRFSFEAEVKGKLRALSILF-----EVGTEENLDFKAIIDGVESVSFRGKQA 210

QY 69 -----TIFNLLDKGSDAIFPVFNGGLFS---EDKVYLNNEGLLSISETE-----EIL- 114
Db 211 ALDPFILLNLLPNSTD--KYIYNGSLTSPCTDVTVDWIVFKDVTVSISESQLAVFCEVLT 268

QY 115 -----VKMLFFPEEKNIKDEKFKVYSR-----LDPKSFGE 143
Db 269 MQQSGYVLMYLDLQNNFREQQY-KFSRQVFSYTGKEEIIHEAACRSEPNVQADPENYTS 327

QY 144 L-----YETLLE-----YDLRIADTTVHRIEDG 167
Db 328 LLVTWPRVVDYTMIEKFAVLVYQQLDGEDQTKHEFLTGD 367

RESULT 14

US-10-295-027-378
; Sequence 378, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynnne, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.

```

; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 378
; LENGTH: 1337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-378

```

```

Query Match      11.1%; Score 97; DB 4; Length 1337;
Best Local Similarity 21.8%; Pred. No. 7.6;
Matches 48; Conservative 36; Mismatches 50; Indels 86; Gaps 11;

Qy 26 YIEDNIP-----KKILOENKLYRSSISFRYPFYDENTKKLEVKKII-----68
Db 38 YCFDADRFSSFEAVKGGKLRALSILF-----EVGTENLDPKAIIDGVESVSRFGKQA 92
Qy 69 -----TIFNLDKGSDAIKFPVFNGLFS---EDKVKYLNNEGLLSISEIE-----EIL- 114
Db 93 ALDPPFILLNLLPNSTD--KYYIYNGSLTSPCTDTVDWIVFKDTVSISESQLAVFCEVLT 150
Qy 115 -----VKMLFFEEKNIKDEKFKVYSR-----LDPKSFGE 143
Db 151 MQQSGYVLMMDYLQNNFRQQY-KFSRQVFSSTGKEIHEAVCSBPENVQADPENYTS 209
Qy 144 L-----YETLLE-----YDLRIADTTVHRIIEDG 167
Db 210 LLVTWRPRVYDTMIKFAVLVYQQLDGEDQTKHEFLTDG 249

```

```

RESULT 15
US-10-295-027-1266
; Sequence 1266, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glyne, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Murray, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.

```

```

; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1266
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1266

Query Match      11.1%; Score 97; DB 4; Length 1359;
Best Local Similarity 21.8%; Pred. No. 7.8;
Matches 48; Conservative 36; Mismatches 50; Indels 86; Gaps 11;

Qy 26 YIEDNIP-----KKILOENKLYRSSISFRYPFYDENTKKLEVKKII-----68
Db 156 YCFDADRFSSFEAVKGGKLRALSILF-----EVGTENLDPKAIIDGVESVSRFGKQA 210
Qy 69 -----TIFNLDKGSDAIKFPVFNGLFS---EDKVKYLNNEGLLSISEIE-----EIL- 114
Db 211 ALDPPFILLNLLPNSTD--KYYIYNGSLTSPCTDTVDWIVFKDTVSISESQLAVFCEVLT 268
Qy 115 -----VKMLFFEEKNIKDEKFKVYSR-----LDPKSFGE 143
Db 269 MQQSGYVLMMDYLQNNFRQQY-KFSRQVFSSTGKEIHEAVCSBPENVQADPENYTS 327
Qy 144 L-----YETLLE-----YDLRIADTTVHRIIEDG 167
Db 328 LLVTWRPRVYDTMIKFAVLVYQQLDGEDQTKHEFLTDG 367

Search completed: January 24, 2006, 20:54:33
Job time : 31.5815 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:59:25 ; Search time 2.78483 Seconds
(without alignments)
633.203 Million cell updates/sec

Title: US-10-688-058-14
Perfect score: 875
Sequence: 1 QITQHILASILESIIFILR.....IADTTVHRIEDGVYLIRTE 174

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 70609 seqs, 10134256 residues

Total number of hits satisfying chosen parameters: 70609

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New.*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US03_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	11.1	374	7 US-11-097-728-4	Sequence 4, Appli
2	97	11.1	2314	7 US-11-097-728-2	Sequence 2, Appli
3	97	11.1	2353	7 US-11-097-728-6	Sequence 6, Appli
4	95.5	10.9	331	6 US-10-793-626-3256	Sequence 3256, Ap
5	93	10.6	389	7 US-11-069-642-18	Sequence 18, Appl
6	90	10.3	482	6 US-10-793-626-24	Sequence 24, Appl
7	89.5	10.2	249	6 US-10-793-626-3298	Sequence 3298, Ap
8	82.5	9.4	2710	6 US-11-051-453-41	Sequence 41, Appl
9	82	9.4	833	6 US-10-467-657-3876	Sequence 3876, Ap
10	80	9.1	222	6 US-10-793-626-128	Sequence 128, App
11	80	9.1	347	6 US-10-793-626-106	Sequence 106, App
12	80	9.1	989	6 US-10-793-626-2594	Sequence 2594, Ap
13	79.5	9.1	371	6 US-10-793-626-3294	Sequence 3294, Ap
14	77	8.8	199	6 US-10-793-626-1722	Sequence 1722, Ap
15	77	8.8	312	7 US-11-156-084-331	Sequence 331, App
16	76.5	8.7	396	7 US-11-022-562-228	Sequence 228, App
17	76	8.7	398	7 US-11-129-143-74	Sequence 74, Appl
18	74.5	8.5	449	6 US-10-485-517-272	Sequence 272, App
19	74	8.5	393	6 US-10-485-517-316	Sequence 316, App
20	74	8.5	437	7 US-11-069-642-92	Sequence 92, Appl
21	74	8.5	438	7 US-11-069-642-47	Sequence 47, Appl
22	74	8.5	438	7 US-11-069-642-49	Sequence 49, Appl
23	74	8.5	438	7 US-11-069-642-53	Sequence 53, Appl
24	74	8.5	438	7 US-11-069-642-55	Sequence 55, Appl
25	74	8.5	438	7 US-11-069-642-57	Sequence 57, Appl

26	74	8.5	438	7 US-11-069-642-59	Sequence 59, Appl
27	74	8.5	438	7 US-11-069-642-61	Sequence 61, Appl
28	73.5	8.4	2367	7 US-11-051-453-42	Sequence 42, Appl
29	73	8.3	438	7 US-11-069-642-51	Sequence 51, Appl
30	73	8.3	438	7 US-11-069-642-63	Sequence 63, Appl
31	73	8.3	701	6 US-10-793-626-710	Sequence 710, App
32	72	8.2	413	7 US-11-032-951-12	Sequence 12, Appl
33	71.5	8.2	600	6 US-10-606-302-3	Sequence 3, Appli
34	71	8.1	289	6 US-10-793-626-880	Sequence 880, App
35	71	8.1	325	6 US-10-793-626-1076	Sequence 1076, Ap
36	70.5	8.1	689	7 US-11-113-424-46	Sequence 46, Appl
37	70	8.0	763	6 US-10-873-528-63	Sequence 63, Appl
38	70	8.0	1992	7 US-11-069-834-58	Sequence 58, Appl
39	70	8.0	2000	7 US-11-069-834-56	Sequence 56, Appl
40	69.5	7.9	178	6 US-10-467-657-2220	Sequence 2220, Ap
41	69.5	7.9	404	6 US-10-793-626-1130	Sequence 1130, Ap
42	69.5	7.9	574	7 US-11-024-959-300	Sequence 300, App
43	69	7.9	258	6 US-10-793-626-440	Sequence 440, App
44	69	7.9	277	7 US-11-156-084-180	Sequence 180, App
45	69	7.9	311	7 US-11-156-084-330	Sequence 330, App

ALIGNMENTS

RESULT 1
US-11-097-728-4
; Sequence 4, Application US/11097728
; Publication No. US20050260132A1
; GENERAL INFORMATION:
; APPLICANT: Erik Foehr
; APPLICANT: Sabine Muller
; APPLICANT: Daniel J. Chin
; APPLICANT: Mirella Gonzalez-Zulueta
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES DIRECTED TO RECEPTOR PROTEIN
; FILE REFERENCE: TYROSINE PHOSPHATASE ZETA
; CURRENT APPLICATION NUMBER: US/11/097,728
; PRIOR FILING DATE: 2005-03-31
; PRIOR APPLICATION NUMBER: 10/652,981
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-097-728-4

Query Match 11.1%; Score 97; DB 7; Length 374;
Best Local Similarity 21.8%; Pred. No. 0.12;
Matches 48; Conservative 36; Mismatches 50; Indels 86; Gaps 11;
QY 26 YIEDNDF-----KKILQENKLYRSSISFRYFFVDENTKKKLEYKKII-----68
DB 156 YCFDADRFSFEAEVKGKLRALSILF-----EVGTENLDFKAIIDGVESVSRFQKQA 210
QY 69 -----TIFNLLDKSGDAIKFPFVNGGLFS---EDKVYLNNEGILTSIEI-----EIL- 114
DB 211 ALDPFILLNLLFNSTD--KYIYNGSLTSPCTDTVDWIVFKDTVSISESQLAVFCEVLT 268
QY 115 -----VKMLFFBEKNIKDEKFKVYSR-----LDPKSFG 143
DB 269 MQQSGVYMLMDYLQNNFREQQY-KFSRQVFSYTGKEIHEAVCSSEPVNQADPENYTS 327
QY 144 L-----YETILE-----YDLRIADTTVHRIEDG 167
DB 328 LLVTWRPRVVYDTMIETKFAVLYQQLDGEDQTKHEFLTGD 367

RESULT 2
US-11-097-728-2
; Sequence 2, Application US/11097728

```

; Publication No. US20050260132A1
; GENERAL INFORMATION:
; APPLICANT: Erik Foehr
; APPLICANT: Sabine Muller
; APPLICANT: Daniel J. Chin
; APPLICANT: Mirella Gonzalez-Zulueta
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES DIRECTED TO RECEPTOR PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASE ZETA
; FILE REFERENCE: AGYT-006CIP
; CURRENT APPLICATION NUMBER: US/11/097,728
; CURRENT FILING DATE: 2005-03-31
; PRIOR APPLICATION NUMBER: 10/652,981
; PRIOR FILING DATE: 2003-08-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2314
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-11-097-728-2

Query Match      11.1%; Score 97; DB 7; Length 2314;
Best Local Similarity 21.8%; Pred. No. 1.2;
Matches 48; Conservative 36; Mismatches 50; Indels 86; Gaps 11;

QY      26 YIEDNDIF-----KKILQENKLYRSSISFRFFYDENTKKKLYKKII-----68
DB      156 YCFDADRFSFEAVKGGKRLALSILF-----EVGTENLDFKALIDGVESVSRFGKQA 210

QY      69 -----TIFNLLDKGSDAIFPVFNGGLFS---EDKVYLNNEGLLSISEIE-----EIL- 114
DB      211 ALDPPFILLNLPNST--KYYIYNGSLTSPCTDTVDWIVFKDVTVSISQSLAVFCEVLT 268

QY      115 -----VKMLFFEEKNIKDEKFKVYSR-----LDPKSFGE 143
DB      269 MQQSGYVWMLDYLNQNFREQY-KFSRQVFSYTGKEEIEHVAVCSEPNVQADPENYTS 327

QY      144 L-----YETLLE-----YDLRIADTTVHRIEDG 167
DB      328 LLVTWERPRVVYDTMIEKFAVLVYQQLDGEDQTKHEFLTGD 367

; RESULT 3
; US-11-097-728-6
; Sequence 6, Application US/11097728
; Publication No. US20050260132A1
; GENERAL INFORMATION:
; APPLICANT: Erik Foehr
; APPLICANT: Sabine Muller
; APPLICANT: Daniel J. Chin
; APPLICANT: Mirella Gonzalez-Zulueta
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES DIRECTED TO RECEPTOR PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASE ZETA
; FILE REFERENCE: AGYT-006CIP
; CURRENT APPLICATION NUMBER: US/11/097,728
; CURRENT FILING DATE: 2005-03-31
; PRIOR APPLICATION NUMBER: 10/652,981
; PRIOR FILING DATE: 2003-08-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 2353
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARSPLIC
; LOCATION: (1)...(2353)
; OTHER INFORMATION: PTP-zeta SM2 23a exon variant
; US-11-097-728-6

Query Match      11.1%; Score 97; DB 7; Length 2353;
Best Local Similarity 21.8%; Pred. No. 1.2;
Matches 48; Conservative 36; Mismatches 50; Indels 86; Gaps 11;

```

```

QY      26 YIEDNDIF-----KKILQENKLYRSSISFRFFYDENTKKKLYKKII-----68
DB      156 YCFDADRFSFEAVKGGKRLALSILF-----EVGTENLDFKALIDGVESVSRFGKQA 210

QY      69 -----TIFNLLDKGSDAIFPVFNGGLFS---EDKVYLNNEGLLSISEIE-----EIL- 114
DB      211 ALDPPFILLNLPNST--KYYIYNGSLTSPCTDTVDWIVFKDVTVSISQSLAVFCEVLT 268

QY      115 -----VKMLFFEEKNIKDEKFKVYSR-----LDPKSFGE 143
DB      269 MQQSGYVWMLDYLNQNFREQY-KFSRQVFSYTGKEEIEHVAVCSEPNVQADPENYTS 327

QY      144 L-----YETLLE-----YDLRIADTTVHRIEDG 167
DB      328 LLVTWERPRVVYDTMIEKFAVLVYQQLDGEDQTKHEFLTGD 367

; RESULT 4
; US-10-793-626-3256
; Sequence 3256, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3256
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
; US-10-793-626-3256

Query Match      10.9%; Score 95.5; DB 6; Length 331;
Best Local Similarity 26.0%; Pred. No. 0.15;
Matches 40; Conservative 24; Mismatches 49; Indels 41; Gaps 8;

QY      26 YIEDNDIFKK---ILOENKLYRSSISFRFFYDENTKKKLYKKIITIFNLLDKGSDAI 81
DB      145 YIMTSDINHKTIEYFKQHKYFNVDANHHFFKQDNIVALSEEGKLV-----LNROGHIM 199

QY      82 KPEPVFNGGLF-----SEDKVKY--LNNEGLLSISEIEEILVKML--FFEEKNI 125
DB      200 ETENGNGGVFKSLKKAGYLDKMQQDHVKYIFLNN-----IDNVLVKVLDPLFAGFTV 251

QY      126 KDEKFKYSRLDPK---SFGELY-----ETLLEY 151
DB      252 TQSKDITSKTIQPKDSESVGRLVNVDCKDTVLEY 285

; RESULT 5
; US-11-069-642-18
; Sequence 18, Application US/11069642
; Publication No. US20050260626A1
; GENERAL INFORMATION:
; APPLICANT: LORENS, JAMES B.
; APPLICANT: PRAY, TODD R.
; APPLICANT: KINSELLA, TODD M.
; APPLICANT: BENNETT, MARK K.
; TITLE OF INVENTION: IN VIVO PRODUCTION OF CYCLIC PEPTIDES FOR
; TITLE OF INVENTION: INHIBITING PROTEIN-PROTEIN INTERACTION
; FILE REFERENCE: RIGL-022CIF3
; CURRENT APPLICATION NUMBER: US/11/069,642
; CURRENT FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: 10/232,758

```

; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: 10/422,536
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 09/800,770
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/187,130
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 168
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 18
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-11-069-642-18

Query Match 10.6%; Score 93; DB 7; Length 389;
Best Local Similarity 19.5%; Pred. No. 0.29;
Matches 43; Conservative 41; Mismatches 76; Indels 60; Gaps 7;

QY 5 HILASILESIIFILRIFFIAYIEDN-----DIFKK----- 35
DB 179 YIEVPLFEDPLLLAFGIVSRIEKSTLKISGRNLELFRKHVGTDESEKAKALDELISK 238
QY 36 -----ILOENKLYRSSISFRFFYDENTKKLEKYKLIITIFNLLDKGSD--AIKFPV 85
DB 239 AKESERYPILEE--LRLGLLFGFTRNELRIEENPTYEVLMEILERIERGSPNLAEKIAV 296
QY 86 FNGGLFSEDKVYKLNNEGLLSISEIEILVKMLFFEEKNKDKFKVYSRLDPKPSFGELY 145
DB 297 LEGRKEENYLRILEEGLNGKUTELGKELLEVRNREFDSKVDYVR-----NIV 349
QY 146 ETL-----LEYDLRIADTT--VHRIEDGVYLIRT 173
DB 350 ENLVLPVKEVERIEYGVYDVTTTETHNFVANGILVHNT 389

RESULT 6
US-10-793-626-24
; Sequence 24, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 24
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-24

Query Match 10.3%; Score 90; DB 6; Length 482;
Best Local Similarity 24.9%; Pred. No. 0.7;
Matches 44; Conservative 33; Mismatches 54; Indels 46; Gaps 10;

QY 26 YIEDNDFIKKLOENKLY-----RSSISFRY-FFYDENTKKLEKYKLI--TIF--NLL 74
DB 194 YLSVFNISKTVYKNNKNHVLFTSDSRANSGNFKFIYEEMLKQDLKLVHSIFKPNIA 253
QY 75 DKGS--DAIKFPVFNGLFSEDKVYKLNNEGLLSISEIEILVKMLFFEEKNKDE---- 128
DB 254 NRRSFIDKLFKFPYFLG-----KSKY-----ILVDDYHPMYKLOFRENQEIQQVQHAV 301
QY 129 ---KFKVYSRLDPK-----SFGELYETLLEYDLRIADT---TVHRIEDGV 168

; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: 10/422,536
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 09/800,770
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/187,130
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 168
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 18
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-11-069-642-18

Query Match 10.6%; Score 93; DB 7; Length 389;
Best Local Similarity 19.5%; Pred. No. 0.29;
Matches 43; Conservative 41; Mismatches 76; Indels 60; Gaps 7;

QY 5 HILASILESIIFILRIFFIAYIEDN-----DIFKK----- 35
DB 179 YIEVPLFEDPLLLAFGIVSRIEKSTLKISGRNLELFRKHVGTDESEKAKALDELISK 238
QY 36 -----ILOENKLYRSSISFRFFYDENTKKLEKYKLIITIFNLLDKGSD--AIKFPV 85
DB 239 AKESERYPILEE--LRLGLLFGFTRNELRIEENPTYEVLMEILERIERGSPNLAEKIAV 296
QY 86 FNGGLFSEDKVYKLNNEGLLSISEIEILVKMLFFEEKNKDKFKVYSRLDPKPSFGELY 145
DB 297 LEGRKEENYLRILEEGLNGKUTELGKELLEVRNREFDSKVDYVR-----NIV 349
QY 146 ETL-----LEYDLRIADTT--VHRIEDGVYLIRT 173
DB 350 ENLVLPVKEVERIEYGVYDVTTTETHNFVANGILVHNT 389

RESULT 6
US-10-793-626-24
; Sequence 24, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 24
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-24

Query Match 10.3%; Score 90; DB 6; Length 482;
Best Local Similarity 24.9%; Pred. No. 0.7;
Matches 44; Conservative 33; Mismatches 54; Indels 46; Gaps 10;

QY 26 YIEDNDFIKKLOENKLY-----RSSISFRY-FFYDENTKKLEKYKLI--TIF--NLL 74
DB 194 YLSVFNISKTVYKNNKNHVLFTSDSRANSGNFKFIYEEMLKQDLKLVHSIFKPNIA 253
QY 75 DKGS--DAIKFPVFNGLFSEDKVYKLNNEGLLSISEIEILVKMLFFEEKNKDE---- 128
DB 254 NRRSFIDKLFKFPYFLG-----KSKY-----ILVDDYHPMYKLOFRENQEIQQVQHAV 301
QY 129 ---KFKVYSRLDPK-----SFGELYETLLEYDLRIADT---TVHRIEDGV 168

DB 302 GAPKTVGFSRTCKKGGPFIDSIGHRNYSKAYSSNNDILYAEAFGIEHRVPTGV 358

RESULT 7
US-10-793-626-3298
; Sequence 3298, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3298
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-3298

Query Match 10.2%; Score 89.5; DB 6; Length 249;
Best Local Similarity 27.3%; Pred. No. 0.34;
Matches 38; Conservative 20; Mismatches 42; Indels 39; Gaps 8;

QY 50 RYFFYDENTKKLEKYKLIITIFN-----LLDKGSDAIKFPVFNGLF-----SE 93
DB 18 KYFNVDANHIFPKQDNIVALSEEGKVLNRDGHIMETNGGVFKSLKAGYLDKMQQ 77
QY 94 DKVKY--LNNEGLLSISEIEIEILVKML--FFEEKNKDEKFKVYSRLDPK---SFGELY- 145
DB 78 DHVKYIFLNN-----IDNVLVKVLDPFLFAGFTVTSKDTISKTIQKDSVGLV 129

QY 146 ----ETLLEY---DLRIAD 157
DB 130 VDCKDTVLEYSELDTDIAN 148

RESULT 8
US-11-051-453-41
; Sequence 41, Application US/11051453
; Publication No. US20050287150A1
; GENERAL INFORMATION:
; APPLICANT: AMBROSINO, DONNA
; APPLICANT: BABCOCK, GREGORY J.
; APPLICANT: BROERING, THERESA
; APPLICANT: GRAZIANO, ROBERT
; APPLICANT: HERNANDEZ, HECTOR JAVIER
; APPLICANT: LOWY, ISRAEL
; APPLICANT: MANDELL, ROBERT
; APPLICANT: MOLRINE, DEBORAH
; APPLICANT: THOMAS, JR., WILLIAM D.
; APPLICANT: ZHANG, HUI-FEN
; TITLE OF INVENTION: ANTIBODIES AGAINST CLOSTRIDIUM DIFFICILE TOXINS AND
; FILE REFERENCE: MJ1-001
; CURRENT APPLICATION NUMBER: US/11/051,453
; CURRENT FILING DATE: 2005-02-04
; PRIOR APPLICATION NUMBER: 60/542,357
; PRIOR FILING DATE: 2004-02-06
; PRIOR APPLICATION NUMBER: 60/613,854
; PRIOR FILING DATE: 2004-09-28
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: Patent in Ver. 3.3
; SEQ ID NO 41
; LENGTH: 2710
; TYPE: PRT
; ORGANISM: Clostridium difficile

US-11-051-453-41

Query Match 9.4%; Score 82.5; DB 7; Length 2710;
 Best Local Similarity 23.4%; Pred. No. 27;
 Matches 37; Conservative 20; Mismatches 34; Indels 67; Gaps 8;
 QY 6 ILASILEES--LIFILRI-----PFIAYIEDNIFPKILOEN 40
 DB 1747 ILVRYLEENKKILQKIRKIGLUNTSQFNKMSIDFKDKLGLVIMSN--FKSPNSN 1804
 QY 41 KLYRSISPRY-----PFYDENTKKLEYKIIITI-----FNL----- 73
 DB 1805 ELDRDLHGFKIDNKYYIYDESKL---VKGLININNSLFYDPFIEFNLVTGWTQINGKK 1861
 QY 74 ----LKGSDAIFPVFNGGLFSEDKVKYLNNEGLLSI 107
 DB 1862 YYPDINTGAALTSYKIINGKH-----YFNNDGVNQL 1893

RESULT 9

US-10-467-657-3876
 ; Sequence 3876, Application US/10467657
 ; Publication No. US20050260581A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIRON Spa
 ; APPLICANT: FONTANA Maria Rita
 ; APPLICANT: PIZZA Mariagrazia
 ; APPLICANT: MASIGNANI Vega
 ; APPLICANT: MONACI Elisabetta
 ; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
 ; FILE REFERENCE:
 ; CURRENT APPLICATION NUMBER: US/10/467,657
 ; CURRENT FILING DATE: 2003-08-11
 ; PRIOR APPLICATION NUMBER: GS-0103424.8
 ; PRIOR FILING DATE: 2001-02-12
 ; NUMBER OF SEQ ID NOS: 9218
 ; SOFTWARE: SeqWin99, version 1.04
 ; SEQ ID NO 3876
 ; LENGTH: 833
 ; TYPE: PRT
 ; ORGANISM: Neisseria gonorrhoeae
 US-10-467-657-3876

Query Match 9.4%; Score 82; DB 6; Length 833;
 Best Local Similarity 20.6%; Pred. No. 6.9;
 Matches 44; Conservative 41; Mismatches 65; Indels 64; Gaps 11;

QY 1 QITQHILASILE-----ESIIFILRIFFIAYIEDN-----DIFPKILOE 39
 DB 439 KFTBELKQKILQKYPDIKPLVNLQFLSDGIIDNENFAEDGYDKLAAYPEAPFKGLDK 498
 QY 40 NKLY-----RSSISPRYFFYDE-----NTKKLEYK-----KIITIFNLLDKGSDA 80
 DB 499 GKVSNADEGKDTIMREGKYEELKALWELHKAVALQYKIKDEAEFVDFLTAYLR-ENA 557
 QY 81 IKFPVNGGLFSEDKVKYLNNEGLLSISEIEILVKMLFPEEKNIKDEKFKVKSRLDPKS 140
 DB 558 AKFP--QAGICTAVNEAYINN-GLMLSRID-----SIEDDFIRENTMTYRE 602
 QY 141 FGELEYETLLELDRIADTTVHRIIEDGVYLRTE 174
 DB 603 FLE-----KLAOTAKIQMOTLHQ-----AFYVRDE 628

RESULT 10

US-10-793-626-128
 ; Sequence 128, Application US/10793626
 ; Publication No. US20050255478A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KIMMERLY, WILLIAM JOHN
 ; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
 ; FILE REFERENCE: PU3480US
 ; CURRENT APPLICATION NUMBER: US/10/793,626

; CURRENT FILING DATE: 2004-03-04
 ; PRIOR APPLICATION NUMBER: 60/164,258
 ; PRIOR FILING DATE: 1999-11-09
 ; NUMBER OF SEQ ID NOS: 4472
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 128
 ; LENGTH: 222
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic
 ; OTHER INFORMATION: amino acid sequence
 US-10-793-626-128

Query Match 9.1%; Score 80; DB 6; Length 222;
 Best Local Similarity 27.0%; Pred. No. 2;
 Matches 41; Conservative 19; Mismatches 52; Indels 40; Gaps 8;
 QY 27 IEDN-DIFPKILOENKLYRSSISFRYFFYDENTKK-----KLEYKIIITIFNLLDKGSDAI 81
 DB 92 VENNITQTEKRFLYERPVR-----YDSVTPELELLKRNVDKNVTFGNLPK--KAI 140
 QY 82 KFPVNGGLFSEDKVKYLNNEGLLSISEIEILVKMLFPEEKNIKDEKFKVKSRLDPKSF 141
 DB 141 QVPKYTGTTTDPFVYMI-----ETDEQDAKYLIVETK-----AENMLGDKSI 184
 QY 142 GEL-----YETL-----LEYDLRIADTTVHRIIE 165
 DB 185 GRIQKKFFNTLDNLNIKYLQATSAQDVYNEIK 216

RESULT 11

US-10-793-626-106
 ; Sequence 106, Application US/10793626
 ; Publication No. US20050255478A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KIMMERLY, WILLIAM JOHN
 ; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
 ; FILE REFERENCE: PU3480US
 ; CURRENT APPLICATION NUMBER: US/10/793,626
 ; CURRENT FILING DATE: 2004-03-04
 ; PRIOR APPLICATION NUMBER: 60/164,258
 ; PRIOR FILING DATE: 1999-11-09
 ; NUMBER OF SEQ ID NOS: 4472
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 106
 ; LENGTH: 347
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic
 ; OTHER INFORMATION: amino acid sequence
 US-10-793-626-106

Query Match 9.1%; Score 80; DB 6; Length 347;
 Best Local Similarity 21.7%; Pred. No. 3.5;
 Matches 35; Conservative 32; Mismatches 56; Indels 38; Gaps 6;
 QY 31 DIFPKILOENKLYRSSISFRYFFYDENTKKLE---YKIIITIFNLLDKGSDAIFPVFN 87
 DB 125 NVAPQIERQEKYQOKMLHRSIQAOENRKRISRELHDSVIQDMLNDELVELLLKYK--- 181
 QY 88 GGLFSEDKVKYLNNEGLLSISEIEILVKMLFPEEKNIKDEKFKVKSRLDP----- 138
 DB 182 ----HRDKV-----LAETSQRIEGLLSQLIDIRNMSVE--LRPSSLODLGIEAAFKS 228
 QY 139 --KSPGLEYETLLEYDLRI-----ADTTVHRIIEDGVY 169
 DB 229 YFQKFEENYGMHIXYDSNIKGMRFDNEIETVYVYVQEGVF 269

RESULT 12

US-10-793-626-2594


```

Qy 125 IKDEKFKVKYSRLDPKSF 141
      :      : | : :
Db 147 LPQLPVISFLRSEYEKY 163

RESULT 14
US-10-793-626-1722
; Sequence 1722, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMWERY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P034800S
; CURRENT APPLICATION NUMBER: US/10793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1722
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-1722

Query Match      8.8%; Score 77; DB 6; Length 199;
Best Local Similarity 26.8%; Pred. No. 3.2; Indels 20; Gaps 5;
Matches 34; Conservative 20; Mismatches 57;

Qy 24 IAYIEDNDIPKKILOENKLYRSSISFRYFYDENTKKCLEYKKIITIFNLLDKGS----- 78
      : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 56 IIVADDNYKWLQILPDKKRYSMVMP-----DNKGNPLEYFDINIKITOKGNARTID 109
      : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 79 ---DAIKFPVNGGLSFSEDKVYLNNEGLLSISEIE--ILVKMLFFE-EKN---IKDEK 129
      : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 110 LCLDLVLFPNGEYELVDEDDLMTALQNKQISKQYHEAVIIAHQLMIEIEDNFSFIQDKV 169
      : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 130 FVKYSRLDPKS 140
      | : : : | :
Db 170 MRCYHKINHA 180

RESULT 15
US-11-156-084-331
; Sequence 331, Application US/11156084
; Publication No. US20060010515A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology LLC
; TITLE OF INVENTION: Controlled expression of cytokinin biosynthetic genes leads to
; TITLE OF INVENTION: agronomically interesting phenotypes
; FILE REFERENCE: (38-21)
; CURRENT APPLICATION NUMBER: US/11/156,084
; CURRENT FILING DATE: 2005-06-17
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 331
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-11-156-084-331

Query Match      8.8%; Score 77; DB 7; Length 312;
Best Local Similarity 24.3%; Pred. No. 5.6;
Matches 28; Conservative 23; Mismatches 46; Indels 18; Gaps 4;

Qy 65 KKIITIFNLLDKGSDAIKFPV---FNGGLFSFDKVKYLNNEGLLSISEIEIILV---KML 118
      : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 12 KKIALLGPGSGKSLSTELAQELDAEIFSLD-----SUSIYEDINIASAKPS 60
      : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 119 FFEKNIKDKFKVKYSRLDPKSFGEYETILLEYDLRIADTTVHRIIEDGVYLIRT 173

```

Db 61 LKERNIK-HYALDYNIDEKKNASLFKILLEDAMRVSSKEILLIVGSSFYLKS 114

Search completed: January 24, 2006, 20:55:37
Job time : 3.78483 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:21:03 ; Search time 285.376 Seconds
(without alignments)
1966.133 Million cell updates/sec

Title: US-10-688-058-16
Perfect score: 6538
Sequence: 1 MKTNDIVKTNPNISLYKQL.....SSLIKGYDNLNNKIPYHK 1277

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_21.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*
- 9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6534	99.9	1277	8	Adu98757 Borrelia
2	872	13.3	174	8	Adu98755 Borrelia
3	445	6.8	1203	8	Adv87926 Streptoc
4	445	6.8	1203	8	Adv81375 Streptoc
5	445	6.8	1203	8	Adv79179 Streptoc
6	435	6.7	1532	4	Aag90567 C. glutami
7	419.5	6.4	1279	2	Aaw98761 H. pylori
8	383	5.9	1009	5	Adu24049 Bacillus
9	382	5.8	1009	5	Aao20949 Protein o
10	382	5.8	1009	8	Adu24050 Bacillus
11	382	5.8	1009	8	Adu24023 Bacillus
12	376	5.8	1009	8	Adu24051 Bacillus
13	371.5	5.7	1644	7	Adc00979 Enterohae
14	365.5	5.6	1282	8	Adn46345 Thermococ
15	364	5.6	1629	6	Aau48711 Propionib
16	364	5.6	1629	6	Aam45230 Propionib
17	349.5	5.3	1979	3	Aab18171 Plasmodiu
18	336.5	5.1	3973	3	Aab18253 Plasmodiu
19	326.5	5.0	2295	3	Aab18180 Plasmodiu
20	316	4.8	1639	2	Aaw54145 P. falcip
21	316	4.8	1639	5	Aae29345 Plasmodiu
22	315	4.8	2184	4	Aae00425 P. falcip
23	312	4.8	1516	3	Aab18195 Plasmodiu
24	309	4.7	1654	1	Aap50777 Sequence

ALIGNMENTS

RESULT 1

Adu98757

ID Adu98757 standard; protein; 1277 AA.

AC Adu98757;

XX Adu98757;

DT 24-FEB-2005 (first entry)

XX 24-FEB-2005 (first entry)

DE Borrelia burgdorferi antigenic polypeptide seqid 16.

XX antibacterial; vaccine; immune stimulation; immunity; antigen;

KW DNA library.

XX Borrelia burgdorferi.

OS Borrelia burgdorferi.

XX WO2004103269-A2.

PN 02-DEC-2004.

XX 17-OCT-2003; 2003WO-US033056.

PF 18-OCT-2002; 2002US-0419401P.

XX (NACR-) MACROGENICS INC.

PA (TEXA) UNIV TEXAS SYSTEM.

XX Sykes KF, Hale KS, Johnston SA;

PI WPI; 2004-834155/82.

XX N-PSDB; Adu98756.

DR Immunizing a subject against Borrelia burgdorferi infection comprises

PT providing to the subject at least one Borrelia antigen or its fragment.

XX Claim 27; SEQ ID NO 16; 121pp; English.

PS The invention describes a method of immunizing a subject comprising

CC providing to the subject at least one Borrelia antigen or its fragment to

CC induce an immune response. Also described are: an isolated polynucleotide

CC comprising a sequence having at least 17 contiguous nucleotides in common

CC with a sequence not given in the specification; an isolated polypeptide

CC having at least 5 consecutive amino acids of the sequence not given in

CC the specification; a vaccine composition comprising at least one Borrelia

CC antigen or at least one polynucleotide encoding a Borrelia antigen;

CC screening for at least one test polypeptide or test polynucleotide

CC encoding a polypeptide for an ability to produce an immune response;

CC preparing a vaccine; vaccinating a subject; treating a subject infected

25 307 4.7 2485 3 AAB18172 Aab18172 Plasmodiu
26 301.5 4.6 1163 8 ABU23940 Abu23940 Protein e
27 301.5 4.6 2024 6 ADP25444 Adp25444 Plasmodiu
28 298 4.6 2013 3 AAB18265 Aab18265 Plasmodiu
29 295.5 4.5 1875 6 ABR53560 Abr53560 Protein s
30 295.5 4.5 1875 7 ADK64380 Adk64380 Disease t
31 295.5 4.5 1875 8 ADS43855 Ads43855 Bacterial
32 293 4.5 1178 6 ABU24510 Abu24510 Protein e
33 293 4.5 1196 6 ABU24813 Abu24813 Protein e
34 293 4.5 2688 4 AAM40883 Aam40883 Human pol
35 292.5 4.5 1817 3 AAB18301 Aab18301 Plasmodiu
36 292 4.5 1308 3 AAB18167 Aab18167 Plasmodiu
37 291.5 4.5 1169 6 ABU19294 Abu19294 Protein e
38 288 4.4 1132 9 ADZ10441 Adz10441 P. gingiv
39 287 4.4 1346 3 AAB18236 Aab18236 Plasmodiu
40 285.5 4.4 2663 4 AAM39097 Aam39097 Human pol
41 285.5 4.4 2663 8 ADQ17932 Adq17932 Human sof
42 285.5 4.4 2663 9 ADX06867 Adx06867 Cyclin-de
43 284 4.3 1000 8 ADS91313 Ads91313 A. calcoa
44 284 4.3 1184 6 ABU25330 Abu25330 Protein e
45 284 4.3 2010 3 AAB18218 Aab18218 Plasmodiu

CC with a pathogen; passive immunization; vaccination; detecting Herpesvirus
CC ; and detecting antibodies to a herpesvirus. The method is useful in
CC immunizing a subject against Borrelia burgdorferi infection. This is the
CC amino acids sequence of a Borrelia burgdorferi polypeptide that may be
CC useful in immunizing a subject against infection by a member of the
CC Borrelia species.
XX
SQ

Sequence 1277 AA;

Query Match 99.9%; Score 6534; DB 8; Length 1277;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKNNDIVKTNPNISLYKQISKDFIKENINKLKDFILINKLFSIDDNSTANTIESLL 60
DB 1 MKNNDIVKTNPNISLYKQISKDFIKENINKLKDFILINKLFSIDDNSTANTIESLL 60
QY 61 KYIFEELNYVEQKAGQIEGVSVDILLFENDKDKASFNKKLKAQKNEPIEDIL 120
DB 61 KYIFEELNYVEQKAGQIEGVSVDILLFENDKDKASFNKKLKAQKNEPIEDIL 120
QY 121 IIAEVKRPFSFPAKDKLKESEDQLYRYLNQYQKHGILSNQWRLYDKSKVLYGEKRY 180
DB 121 IIAEVKRPFSFPAKDKLKESEDQLYRYLNQYQKHGILSNQWRLYDKSKVLYGEKRY 180
QY 181 IEFNFKXIEEBEYKEQEWFLFYLRKERYLKTNSVIEVEKEQIAKEEIIQKTLKEI 240
DB 181 IEFNFKXIEEBEYKEQEWFLFYLRKERYLKTNSVIEVEKEQIAKEEIIQKTLKEI 240
QY 241 LYERPDDSIYVFKAKNIYDEFKLSGKEITQHILASLLESIIFILRIPIAYIENDIF 300
DB 241 LYERPDDSIYVFKAKNIYDEFKLSGKEITQHILASLLESIIFILRIPIAYIENDIF 300
QY 301 KKILQENKLYRSSISFRFFYDENTKKLEYKKIITIFNLLDKGSDAIFPPVFNGLFSE 360
DB 301 KKILQENKLYRSSISFRFFYDENTKKLEYKKIITIFNLLDKGSDAIFPPVFNGLFSE 360
QY 361 DKVKYLNNEGLLSISEEILVKMLPFEENKIODEKVKYSRLDPKSFGELYETLLEYDL 420
DB 361 DKVKYLNNEGLLSISEEILVKMLPFEENKIODEKVKYSRLDPKSFGELYETLLEYDL 420
QY 421 RIADTVHRIIEBGVYLIRTEEELENKKVANKIATYLGNIYLTSRSLDRKKSAGYTPDD 480
DB 421 RIADTVHRIIEBGVYLIRTEEELENKKVANKIATYLGNIYLTSRSLDRKKSAGYTPDD 480
QY 481 LTDFMWISSIEEQLKTKSPLDIKIIDNSCGSGHPLISCLDYLTEKWYELDKFEDVKEL 540
DB 481 LTDFMWISSIEEQLKTKSPLDIKIIDNSCGSGHPLISCLDYLTEKWYELDKFEDVKEL 540
QY 541 DEEYRVIIIESEEYDVQDSISKELVLRMLLKXCIYGVNDINPISVEITMLSLWINTPIFG 600
DB 541 DEEYRVIIIESEEYDVQDSISKELVLRMLLKXCIYGVNDINPISVEITMLSLWINTPIFG 600
QY 601 TPLSFIEHHIKAGNALLGYTKDEPFDIVKKKFSFGPSLFFKKRIKEIITILEDIYQKIKGI 660
DB 601 TPLSFIEHHIKAGNALLGYTKDEPFDIVKKKFSFGPSLFFKKRIKEIITILEDIYQKIKGI 660
QY 661 NDTTKEDIEKSKYIKEYEESKDIDNLRIFSLIKYSLSPDKSLNMEFSDIASVISLIE 720
DB 661 NDTTKEDIEKSKYIKEYEESKDIDNLRIFSLIKYSLSPDKSLNMEFSDIASVISLIE 720
QY 721 NILGNKTSSEDEKIEKIRKLSSYYKGFHYGIEPDPDIOEGPDIVGNPPWEKTKFNTEF 780
DB 721 NILGNKTSSEDEKIEKIRKLSSYYKGFHYGIEPDPDIOEGPDIVGNPPWEKTKFNTEF 780
QY 781 FSKHIPNRYKLGIEQNIIEKQIILSKONHPLSIEVNEEKNSIIANNIYKDFKCFYSGG 840
DB 781 FSKHIPNRYKLGIEQNIIEKQIILSKONHPLSIEVNEEKNSIIANNIYKDFKCFYSGG 840
QY 841 DPNLFRYFVTNFKLIEKEGNLTYLVPSAIWNNESSRILRHIFARYKLANIYGFENKRR 900
DB 841 DPNLFRYFVTNFKLIEKEGNLTYLVPSAIWNNESSRILRHIFARYKLANIYGFENKRR 900

QY 901 FKDVHSSFKFAIFQLSNIKESSTSSFKAKFMIOSSNLIKETRLDKSDDAYKGIELNI 960
DB 901 FKDVHSSFKFAIFQLSNIKESSTSSFKAKFMIOSSNLIKETRLDKSDDAYKGIELNI 960
QY 961 NQIKKLSPIQESIIIEFKONEEFTLINKMPKSFALSALGEGYIDFKKGLDPSIKNRKSLKKEC 1020
DB 961 NQIKKLSPIQESIIIEFKONEEFTLINKMPKSFALSALGEGYIDFKKGLDPSIKNRKSLKKEC 1020
QY 1021 NKNKLIFFLYSGANIHQPNRSRFFEDKAKESSKLLMTDKEDLEKVLTKDNOYQOTERVFYRA 1080
DB 1021 NKNKLIFFLYSGANIHQPNRSRFFEDKAKESSKLLMTDKEDLEKVLTKDNOYQOTERVFYRA 1080
QY 1081 IASNTNERTMISTLSPGNCYCVNSIYINDEKTPISLYKKLFIISFNSFVDFLLRRFVD 1140
DB 1081 IASNTNERTMISTLSPGNCYCVNSIYINDEKTPISLYKKLFIISFNSFVDFLLRRFVD 1140
QY 1141 SNVLKSCLYQCMPQPEEKEILSNPLYNLAKNTSLIIAKNDPENPKYLLYLEYFKFDKE 1200
DB 1141 SNVLKSCLYQCMPQPEEKEILSNPLYNLAKNTSLIIAKNDPENPKYLLYLEYFKFDKE 1200
QY 1201 KVNKILKLDKEDEFFKEKENENNFIASLYSLAKEDFITLLGDFKALKNKKKGEDYISSL 1260
DB 1201 KVNKILKLDKEDEFFKEKENENNFIASLYSLAKEDFITLLGDFKALKNKKKGEDYISSL 1260
QY 1261 IKGYDNYLLNNKIFVHK 1277
DB 1261 IKGYDNYLLNNKIFVHK 1277

RESULT 2

ADU98755
ID ADU98755 standard; protein; 174 AA.
XX
AC ADU98755;
XX AC
DT 24-FEB-2005 (first entry)
XX
DE Borrelia burgdorferi antigenic polypeptide seqid 14.
XX
KW antibacterial; vaccine; immune stimulation; immunity; antigen;
KW DNA library.
OS Borrelia burgdorferi.
XX
XX WO2004103269-A2.
PN
PD 02-DEC-2004.
PP 17-OCT-2003; 2003WO-US033056.
XX
XX 18-OCT-2002; 2002US-0419401P.
PR
XX (MACR-) MACROGENICS INC.
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Sykes KP, Hale KS, Johnston SA;
XX
DR WPI; 2004-834155/82.
DR N-PSDB; ADU98754.
XX
PT Immunizing a subject against Borrelia burgdorferi infection comprises
PT providing to the subject at least one Borrelia antigen or its fragment.
XX
PS Claim 27; SEQ ID NO 14; 121pp; English.
XX

CC The invention describes a method of immunizing a subject comprising
CC providing to the subject at least one Borrelia antigen or its fragment to
CC induce an immune response. Also described are: an isolated polynucleotide
CC comprising a sequence having at least 17 contiguous nucleotides in common
CC with a sequence not given in the specification; an isolated polypeptide
CC having at least 5 consecutive amino acids of the sequence not given in
CC the specification; a vaccine composition comprising at least one Borrelia
CC antigen or at least one polynucleotide encoding a Borrelia antigen;

CC screening for at least one test polypeptide or test polynucleotide
 CC encoding a polypeptide for an ability to produce an immune response;
 CC preparing a vaccine; vaccinating a subject; treating a subject infected
 CC with a pathogen; passive immunization; vaccination; detecting Herpesvirus
 CC ; and detecting antibodies to a herpesvirus. The method is useful in
 CC immunizing a subject against Borrelia burgdorferi infection. This is the
 CC amino acids sequence of a Borrelia burgdorferi polypeptide that may be
 CC useful in immunizing a subject against infection by a member of the
 CC Borrelia species.

XX
 SQ Sequence 174 AA;

Query Match 13.3%; Score 872; DB 8; Length 174;
 Best Local Similarity 99.4%; Pred. No. 6.7e-43;
 Matches 173; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 268 EITQHILASILESIIFILRIFFIAYIEDNDIFKKILQENKLYRSSISFRFFYDENTKK 327
 DB 1 QITQHILASILESIIFILRIFFIAYIEDNDIFKKILQENKLYRSSISFRFFYDENTKK 60
 QY 328 KLEYKKIITFNLLDKGSDAIFPPVNGGLFSDKVKYLNNEGILLSISIEEILVQKLPF 387
 DB 61 KLEYKKIITFNLLDKGSDAIFPPVNGGLFSDKVKYLNNEGILLSISIEEILVQKLPF 120
 QY 388 EKNKIDKRFVKYSRLDPKSFGELYETLLLEYDLRIADTTVHRIEDGVYLIRTE 441
 DB 121 EKNKIDKRFVKYSRLDPKSFGELYETLLLEYDLRIADTTVHRIEDGVYLIRTE 174

RESULT 3

ADV87926
 ID ADV87926 standard; protein; 1203 AA.

AC
 AC ADV87926;

XX
 XX 24-FEB-2005 (first entry)

XX Streptococcus agalactiae protein sequence, SEQ ID 320.

XX Antibacterial; Vaccine; bacterial infection.

XX Streptococcus agalactiae.

XX FR2824074-A1.

XX 31-OCT-2002.

XX 26-APR-2001; 2001FR-00005642.

XX 26-APR-2001; 2001FR-00005642.

XX (INSP) INST PASTEUR.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX Glaser P, Rusniok C, Chevalier F, Frangeul L, Lalloui L;

XX Zouine M, Couve E, Buchrieser C, Poyart C, Trieu CP, Kunst P;

XX WPI; 2004-101891/11.

XX Genomic nucleotide sequences encoding polypeptides of Streptococcus

XX agalactiae for the development of vaccines, diagnostic tools, DNA chips

XX and identification of therapeutic targets.

XX Claim 6; SEQ ID NO 320; 2687pp; French.

XX The present invention relates to novel Streptococcus agalactiae

XX nucleotide sequences (I; ADV87607-ADV87745) and novel polypeptides (II;

XX ADV87746-ADV89950). The nucleotide sequences encode polypeptides of S.

XX agalactiae involved in the synthesis of amino acids, cell membranes,

XX intermediate (central) metabolism, energetic metabolism, fatty acid and

XX phospholipid metabolism, nucleotide metabolism including purines,

XX pyrimidines and/or nucleosides, regulatory functions, replication,

XX transcription, translation, protein transport, adaptation to atypical

CC conditions, sensitivity to medicines and/or analogues, functions related
 CC to transporters, biosynthesis of cofactors, prosthetic groups and
 CC transporters, cell membrane proteins and cellular machinery. (I) are
 CC useful for the detection and/or amplification of nucleic acids.
 CC Pharmaceutical composition comprising (I) or (II) are useful for
 CC treatment of a bacterial S. agalactiae infection. Note: W0200292818A2 is
 CC equivalent for the present basic patent FR2824074A1. W0200292818A2
 CC contains 6617 sequence whereas the present patent only contains 2344
 CC sequences.

XX
 SQ Sequence 1203 AA;

Query Match 6.8%; Score 445; DB 8; Length 1203;
 Best Local Similarity 19.4%; Pred. No. 4.3e-17;
 Matches 285; Conservative 241; Mismatches 400; Indels 546; Gaps 71;

QY 26 KKENIN-KLKDFPILIKNNKLSIDNSTEAN-----IESLKYIFEEINTSVVEQKAG 77
 DB 25 EKEVFNQKLVSYLTHLKD-----EDESSEYQKLLKVFLESVLPNF--INTS----- 71
 QY 78 QIEGVESRDYILLFENDKDKASFNKKLKEAK---KNEPIPIEDILIIAEVKRPTFSFDA 134
 DB 72 -----SRDLAIY-NGKDANSSGLIFPECKSLFNKSEMMSTEKI----- 109
 QY 135 KDKLKESEDLRYLNO--YQKHG-----ILSGKVWRLYDK---SKVLYGEKRYL---- 181
 DB 110 --NSKAFQEIYVYVYIQLERLFNKNLEIKKCIITNGLSWFVIEAKEPEKHFKNKLDVLT 167
 QY 182 -----EENFXKIEEKEEYKEQEFVLYLIRKERYLKTNSNIEVEKEQIAKE 229
 DB 168 KFRNQLSSNKTDLYSEVIAPEIDKAFEGIVIAHFDSQALVKTSKSIEIKNNLTQL 227
 QY 230 KEIIQKTLKEILYERPDSDIVF---KIAKNYDKFEFLSGKEITOHILASILE----- 279
 DB 228 YRFF--TAENLANKE-----IFTDSNKNKNFYDELLYLMGLEETSGTSKIIISRLKPIK 280
 QY 280 -----ESIIFILRIFFIAYIEDNDI-----FKILOEN-KLYRSSISFRY 318
 DB 281 RQYSFVENIINKLEMKOVSKQEDIAIQLTVMNTNRIILFKLLLESQVLFNKDESRYF 340
 QY 319 FFVDENTKKLEVKKIITIF-----NLLDKGSDAI--KF---PVFNGGLFSEDKVKYLVAN 368
 DB 341 LTYE---KLPNFEELYGLFFAVLAKKVSERNDRVQEKFGYVYPLNSSLFEETEIEB-ISR 395
 QY 369 EGL-----LSISBIEEILVKMLFFEBKNIK--DEK-----FVKYSRLDPKSGELYETL 415
 DB 396 DGIGIDRLPEGDIE-----IFSKTALAGVDKRRKGNINFIY-----LFEFL 438
 QY 416 LEYDLRIADTTVHRIEDGVYLIRTEELENNKKNKIATYLGKNIYLTSRSLDRKSGAY 475
 DB 439 DSYDFSTS-----ISHHEKSKNDLIN--ASVL-GLIF---EKINGYRDGSGF 478
 QY 476 YTPDDLTDFM-----VISSIEQLTKTSPL-DIKI 504
 DB 479 YTPGNITWYMSKAIKRTAAVDKVNELLGNCTVEIEKFAIGHSVENARKVSOAIDLLKV 538
 QY 505 IDNSCGSGHFLISCL-----DYLTEKVMYVELDKFEDVKKELEDEE--- 543
 DB 539 CDPVAGSGHFLVLSILNEITALKSDNILEFDAGNIGNLIQCVINDELVIQDMSGNLF 598
 QY 544 YRVIIIESEYDQDSISKELVKRLMKLKKCYIGVDINPISVEITWLSLWI-----NTTIF 599
 DB 599 YQTGNQOQSEQ--IQKAIFNQ---KRHILENSLFAVDINPSSVNICRLRLWIELLKSSYYV 653
 QY 600 G-----TPLSFIEHHIKAGNALLGYTKDEF---FDIVKKPESGSLFK----- 640
 DB 654 QDITDNTQPVLTLPINIDINIKVGDSSL--HKFEFDYFDMRKTDKDYLSLVKDYKETNN 711
 QY 641 KRKEIITILEIDYOKINGINDTTKEDIKSKYKVEESKDIDLN--RIIFSLIKLYSL 699
 DB 712 KKKV-----ADIEWEK-----IERLKHSDTDTASSPELKLSLIERNKAGAI 754
 QY 700 SFDKSLNMEFDSIASVISLIENILGNKTSSEDEKIEKIRKLSYYPFFHYGIEFPDIOE 759


```

539  CDPAVGSGHFLVSLNBIILAKSDNLILPADGNVYIGNLQCTVYINDELVIQDMGNNFL 598
544  YRVIIIESEEDVQDSISKELVLKRMLLKXCIYGVINPINSVEITMLSLMI---NTFIF 599
599  YOTGNQOSEQ--IQKAFNQ---KRHLENSLPAVDINPSSVNICRLRLWIELLKSSYY 653
600  G-----TPLSFIEHHIKAGNALLGYTKDEF---FDIVKKKFGSGPSLFK----- 640
654  QDTDNQPVLTLPNIDINIKVGDLSL--HKPEFDYEDMRKTDKDYLSLVKDYKETNN 711
641  KRIKEIITILEDIYQKIKGINDTTKEDIKSKIKYKEYEBSKOIDNL--RIIFSILKLYSL 699
712  KRVK-----AEIWEK-----IERLKSFDOTASSPELKLUSRLISERNKAGAI 754
700  SFDKSLNMFSDIASVISLNIENILGNKTSSEDEKIEKIRKLSSYFFHYGIEFPDIOE 759
755  SLFEDENKNSKVAEL-----NKLAKAEKLEYRLKNPLFAKGMEMWREFFBILG 805
760  -----GFDIVIGNPPWEKTYNETEFPFSKHIPNYRKLGIKEQNI IQOELSKDNHPLSI 813
806  EAGEFIGFDLVIGNPPY---IFARNQSTDEMSY----- 837
814  EYNEEKSIITAINNIYKDFKCFSTSGDDPNLFRYVTFNFKLKEKGNLYLVPSAIWNE 873
838  -----YKRTYQ---VSEYQANTYITFMELAYQLLRKGGTFSYIIPN----- 875
874  SSSRLAKHIFARYKLVANIVOFENKKRFDVHSSFKAIFQLSNIKESTSSFRAKFMIO 933
876  -----NFL-----TIQT 882
934  SDNILKEITRDLKSDKDAYGIELNINOIKKL---SPIQESIIEFKDNE----- 980
883  NSRVROQITQSTSD-----VVLNSLDKIFADASVDNCIIFFKKGNPNWIEVAELY 933
981  --EFTLINKM---FSKPSALGEGYIDFKKGLD-----PSIKNRKSLK--E 1019
934  HGEPTNTRVAPDFFGEBIPTFISIMWKYRQAIDIFWKVNHFFPLGRPEISTVKSGIKAYE 993
1020  CNNKQLILFLSGANIHOFNSRFPEDKAKSSKLLWIDKEDLEKV--LTKNOY----- 1071
994  TGKGN-----PKMTAQDKNNRIYHARSRDSDSVRPYLDGENVSRYKLTWNGEYVKYGENL 1048
1072  ---QTERVFYRAIASNTNERTMTISLSPGNCYCNSIYINDEKTPISLYKKLFTISIFNS 1128
1049  AAPRDKKEIFEK-----PRILRQIIPSKSYAVEAVYTSD-----VINDLNS 1090
1129  FVPDFLLRRFVDSNVLSKCLYQCMPQPEEKEILSNPLYNLAKNTSLLIAKNDPENFKY 1188
1091  MVI-----TDIQVNFYLLGLIINSRLI-----SL 1114
1189  LLYLEYKPFDKCE-----KVNKI-----LKLKDEDEFFKE 1217
1115  WFFMKDFKFORRLFPQKVNELGDFPIPDATDSQOEDIAQVVEQLMEEMKQSDPD---TE 1171
1218  KENENNPF-----IIASLYSLAKEDFITLLGDPK 1245
1172  KVNQLNLEIDFLVMDFLDLAEEBKOT--VRDPE 1202

```

RESULTS

RESULT 5	
ADV79179	
ID	ADV79179 standard; protein; 1203 AA.
XX	
AC	ADV79179;
XX	
DT	24-FEB-2005 (first entry)
XX	
DE	Streptococcus agalactiae protein, SEQ ID 320.
XX	
KW	Antibacterial; vaccine; bacterial infection.
XX	
OS	Streptococcus agalactiae.
XX	

Db 68 LNNENNE---VEVITIEFKALKNPNEFIKGGDLNVKAFHESLLSYLTERKEGNNLKLH-IL 123
 Qy 160 SNQVWRLDYDSKV-LYGEKRYIEFNFEXKIEEKEEYKEQENFVLFYI-----LIRKERY 212
 Db 124 ATIKELYIIDANEFVFNKDKIEENAFKNCHDR---KGNDRYKAFYDACQKRLNEBFRS 180
 Qy 213 LKTSNVEVEKEQIAKEKEIIQTKLEILYERPDSDIVFKIAKNIYDKPEFKLSGKE----268
 Db 181 LK-YHYIPLKENLAL---IYQALSFPNFKLIKPYSDANTLNKDFEELLYIILGLEQND 236
 Qy 269 -----ITOHLASIL-----EESIIFIL-----RIFFIAYIEDNIFKILQE 306
 Db 237 KGKILIKPSRTQNSLDAKKEYKNLDDREVMALLIANNRILFLRLLES-----286
 Qy 307 NKLYRSSISPRYFYDEN-----TKKLEYKKIITIP--NLLDKGSDAI-----KF 350
 Db 287 -----LLISFKHF---ENPFLTTENFENFDNLNTLFFEVLAACKNSERLPEIKEDKILEKI 338
 Qy 351 PVFNGGLFSE-----DKVKYLNNEGLLSISBEIEILVKMLPFEEKNIKDEK---FVKY 400
 Db 339 PYLNSLFDKTPLELKGHEIKLNDKKL-----EYKNSVLKXKQYQKEKPLPLKY 391
 Qy 401 SRLDPKSGFGLYETLLEYDRIADTTVHRIEDGVVLIIRTEEBLENKKNKIATYLGNI 460
 Db 392 -----LKFRLRYKE---TTTTPKDKONT--DTSESRLINPSV-----L 425
 Qy 461 YLTSRSLDRKSGAYTPDDLTDFWVSSIE-----BQLKT-----496
 Db 426 GLVFEKLGNGYKESFYTPSITSYMCKESITPIVLDKFNAYQWDCENLKALGEIDRNF 485
 Qy 497 -----KSPDLIKIIDNSCGSGHFLISCLDYLTEKWKWYELDKEP-----DVKKELDE 542
 Db 486 SNEKAKYMLTLRLICDPAVSGHFLVSALEMWV-RVAYELGLTASLYRYDLKLENDE 544
 Qy 543 -----BYRVIIIESEBYDQDSISKELV-LKRMLLKXCIYGVNDINPISVEITMLSL 592
 Db 545 IIIHTPTGEIFNYKPDSENDPHHIOKELFNLKKSIENCLFGVDINPNSCEITKRL 604
 Qy 593 WI-----NTPFEGTPLSFIEHHIKAGNAL-----LGY 619
 Db 605 WIELLYSYVYIFKQKNTALET-LPNIDINIKANSLSRFAKDKKALKSEKRNKLEY 663
 Qy 620 TKDEFDPDVKKKPSGSLFK-KRIKEIIT-----ILEDIYQKI-----657
 Db 664 SIABYKELVK-----IYKDPKILETLTHPIKOSNAVKYAKERYQELKQNPKNDFK 715
 Qy 658 KGINDTTKEDIKSKTIYK---EYBESKDIDNRIIFSLLIKYLSLSPDKSLN---MEFSDI 712
 Db 716 KALNDR---IEKIKAFKLTLPNPPKELFKFKLKEHLELYGKSIIEEANYNGLELEAL 771
 Qy 713 ASVLSLLENILGNKTSSEDEKIEKIRKLSYYKF-----PHYGIEF 754
 Db 772 A-LEKOMANLFFDYPYRPKLDSKVVGLEHFNRYLTSYKDLQDENERYANALEWRFEF 830
 Qy 755 PDIOE-----GPDIVGNPPWEKTKNETEFPKSHIPNYRKLGKIEQNIIOBILSKDN 808
 Db 831 PEVLDDGSGDFDCIIGNPPY-----IRQSHI-KDL 861
 Qy 809 HP-LSIEYNEBKNSIIANNIYKDFKTCFTSGGDPNLFYFVFNMLKLIKKEGNLTYLVP 867
 Db 862 KPLLEKYQDPYNS-----TADIYTYFFALAFHLLKKEGFSAFITS 902
 Qy 868 SAINWNESSRILRKHIFAR-----YKLNYYQFENKPKFQDVHSSFKPAIFQLSNIKES 921
 Db 903 NKYTRAKYGAKLREWLKTKTIVSYMBELNALKVPESAA---VDTST-----IHFIKQT 952
 Qy 922 TSSPKAFMIQSSDNLKIEITRDLKDSKDDAYKGIELNINQIKLSPIQSIITEFKDNEE 981
 Db 953 PSK-----ESEFKYETPTNDKDLKSTPHLMKQNVLSTESFIFANATYLLDLRD---1002
 Qy 982 FTLINKMFSKPSALGEGYIDPKGLDPSIKNRKSLKKECNKNKLIIFYSGANIHQFNSRF 1041
 Db 1003 -----KTESVGTPLKDWIDIQINYGIK-----TGA-----NEAF 1030

Qy 1042 FEDKAKESSKLLWIDKEDLEKVLTKDNOYQOTERVFY-----RAIASNTNERTMTSLSP 1096
 Db 1031 IPTTEKRE-----EILNACKTOEERETERLIKPILRGKDIKRYSYEWAHLWVINT 1081
 Qy 1097 GNCYCVNSIYINDEKTPISLYKGLFIISFNFSFVDFLLRRFVDSNV---LKSCLYQCPM 1153
 Db 1082 HNGYTSS---LKSkipPIDIEKYPAIKAHLDAAH-YDTIATRCDOGDTPVHLRNCAYL---1134
 Qy 1154 POPEKE-----ILSNPLYLNLAKE-----NTSLLI-----1178
 Db 1135 -EDFEKEKIWASGVGEYCMIPGLLILDTNFFEVSKFEGNTKNYLLGLLSKLLTFWLK 1193
 Qy 1179 AKNDP-----ENPKY-LLYLEYKPF---DKEKNVKILKLDKDEFFKEK-----1218
 Db 1194 AKGTPLGDMGAVRNTKINMELPWKITAKNKIADKIILAVDKILOAKEKPKANTQKL 1253
 Qy 1219 ENENNFIIASLYSLAKED 1236
 Db 1254 EXEIDALVYQLYHLTDDE 1271

RESULT 8
 ADU24049
 ID ADU24049 standard; protein; 1009 AA.
 AC ADU24049;
 XX
 DT 27-JAN-2005 (first entry)
 DE Bacillus pumilus BpmI restriction endonuclease mutant protein (D74A).
 KW Type IIG restriction endonuclease; molecular cloning;
 KW gene characterisation; BpmI restriction endonuclease; BpmIRM; mutant;
 KW muten; enzyme.
 XX
 OS Bacillus pumilus.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 74 /note= "Wild type Asp is substituted by Ala"
 FT
 XX US2004219584-A1.
 XX
 PD 04-NOV-2004.
 XX
 XX 15-MAR-2004; 2004US-00800946.
 XX
 XX 20-OCT-2000; 2000US-00693146.
 PR 17-MAY-2002; 2002US-00150028.
 XX
 XX (XUSS/) XU S.
 PA (KOB/) KOBBE D.
 PA (ZHUZ/) ZHU Z.
 PA (SAMU/) SAMUELSON J.
 XX
 PI Xu S, Kobbe D, Zhu Z, Samuelson J;
 XX WPI; 2004-783790/77.
 XX
 PT Altering specificity of Type IIG restriction endonuclease by transforming
 PT cell with DNA encoding catalytic domain and methylase N-terminal of first
 PT endonuclease ligated to methylase C-terminal and specificity domain of
 PT second endonuclease.
 XX
 XX Example 5; Page; 34pp; English.
 PS
 XX The present invention relates to a method of altering the cleavage
 CC specificity of a type IIG restriction endonuclease (E) having cleavage,
 CC methylase (M) and specificity domains. The method involves transforming
 CC cell with DNA encoding catalytic domain and methylase N-terminal of first
 CC endonuclease ligated to methylase C-terminal and specificity domain of

PT endonuclease ligated to methylase C-terminal and specificity domain of
 XX second endonuclease.

Example 5; Page; 34pp; English.

CC The present invention relates to a method of altering the cleavage
 CC specificity of a Type IIG restriction endonuclease (E) having cleavage,
 CC methylase (M) and specificity domains. The method involves transforming
 CC cell with DNA encoding catalytic domain and methylase N-terminal of first
 CC endonuclease ligated to methylase C-terminal and specificity domain of
 CC second endonuclease. The method is useful for cleaving DNA molecules into
 CC small fragments for molecular cloning and gene characterisation and for
 CC overproducing restriction endonucleases by genetic manipulation. The
 CC present sequence is the *Bacillus pumilus* BpmI restriction endonuclease
 CC mutant protein (E1007G). Note: This sequence is not shown in the
 CC specification but is derived from *Bacillus pumilus* BpmI restriction
 CC endonuclease wild type protein shown as SEQ ID NO:4 in page 16-18 of the
 CC specification.

XX Sequence 1009 AA;

Query Match 5.8%; Score 382; DB 8; Length 1009;
 Best Local Similarity 19.4%; Pred. No. 1.6e-13;
 Matches 261; Conservative 199; Mismatches 417; Indels 468; Gaps 55;

QY 14 ISYKQSKDFIKK-NINKLKDFI--LKNKLFSDIDNSTEANTIESLLKIYFELNYS 70
 DB 7 VDKYKAHRSFTLPTYNQETLRNDFIDPLKSLGWDVNTKGTHTI---LRDVIQEEYIE 63

QY 71 VEQQKAGIIGSVSRVDILLFENDKOKAFNNKLEAKKNNEPIPTEDILIIAEVKRPTF 130
 DB 64 IKD-----EETKGNPDYTLINGTRK-----LFVEVKKPSP 94

QY 131 SFDADKDLKESD--QLYRLNVOQKHGILSKGVWRLYDKSKVLYGKRYIEFNFXXI 188
 DB 95 NI-----LKSAAAFQTRRGWSANLGI SVLTWFEHLVIYD-----CRY-----T 134

QY 189 EEKEEYKEQWFLFYILIRKRYLTSNVIEKEQIAKEKIIQKTLKEILYERPDSS 248
 DB 135 PDKSDNEHIAKYVFSYEAEAFDEIKIISVESANSALGDMFDMVNTN--VGETPDEY 192

QY 249 IVFKIAKNIYDKFELSGKEITOHILASILEESIIFIL-----RIPFIAYIENDIFKKI 303
 DB 193 FLOQI--ENWREKLAKTAIKNTE-----LGEEDVNFVQRLNRIIFLRVCEDETIKY- 245

QY 304 LOENKLYRSSISFRYPFYDENTKKLEYKIIITENLLDKGSDAIKFPVFNGLFSEDKV 363
 DB 246 -----ETIKSIKNYEEKLDFQKSDR-----KFNGLFD----- 274

QY 364 KYLNNEGLLSISIEELIVRM--LFPPEKNIKDEKFKYSRLDPKSGELYETLLLEYDL 420
 DB 275 -FIDDTLLLEVEIDSNVLIIFSDLYPPQSP-----YDFSVVDPTLSQIYERFLQEI 327

QY 421 RIADTVHRIIEGVYLIIRTEELEENKKNKIATYLNKNIYLSRSLDRKKSAGYTPDD 480
 DB 328 -----IIESG-----GTFHITESPEVAASNGVVPFKI 355

QY 481 LTDPMV---ISSIEEQKTKSPDIKIDNSCGSGHPLISCLDYLTEKVMYELDKEDVK 537
 DB 356 IVEQIVKDTUTPLTEGKKFKNELCNLKIADICCGSGTPLISSYDFLVEKV-----ME 406

QY 538 KELDEEYRVIIERSE-EYDVQDSISKELVLRMLLKKCIYGVINDIPNISVEITMLSLWINT 596
 DB 407 KIIIEEN---IDDSLVAYEEEGLLTLKAKNILENNLFGVDVNPVAVEAEFSL----- 458

QY 597 FIFGTPUSFIEHHIKAGNALLGYTKDEFFDIVKKPFESGSLFKRIKEIITILEDIYQK 656
 DB 459 -----LKLLE----- 464

QY 657 IKGINDTTKEDIKSKKIYKEVESKDIDNRIIFSLIKLYSLSPDKSLNWFSDIASVI 716
 DB 465 --GENEASVNN-----FIHEHE-----DKIL-----PDLTSLII 490

QY 717 SLIENTLGNK-----TSSDEKEKIEKIRKLSSYKFFHYGIEFPDIOE--GFDIVIGNP 768
 DB 491 KCCNSLVNDKFFEFMPESLEDEILFKANP-----FEWEEFEDIMANGGFDIAGNP 543
 QY 769 PWEKTFNETEFESKHIPNYRKLGIKEQNI IQEILLSKDNHPLSIEYNEEKNII IAINNI 828
 DB 544 PY-----VRIQNM-----KKYSPEEIEYYQSKDSE----- 568
 QY 829 KYDFDKFTSGGDPNLFYFVTFN--LKLKEKGNITLYVPSAIWNNESSRILRKHFAR 886
 DB 569 -----YTVAKKETVDKYLFIERALILINPTGLGYIIPHKFFITKGGKELKRTIAEK 621
 QY 887 YKLVNIYQENKGRKPDVHSSKFAIPOLSNIKESISSKAKFMIOSSNLLKEITRDLX 946
 DB 622 HQISKIINFGVTQVFPGRATYTAILLIQAKM-----AQPKYKKNISAEITL----- 669
 QY 947 DSKDD--AVYKGLTELNQIKKLSPIQESIIIEFKDNEEFTLINMKFSGFALGSGYIDFKK 1004
 DB 670 DSEENTCVYSSEKYNSDPMIFLSPETEAV-----FTKFTB--AQFEKLGJ-ITDISV 718
 QY 1005 GL-----DPSIKNRKSLLE-----C-----NNKNLI 1026
 DB 719 GLQTSADKIYIFIPENETSDTVIFNYKGRKYEIEKSIICCPAIVDLSFGSPESIQGNAQMI 778
 QY 1027 FLYSGANIHQFNSRPFEDKAKESSKLWDKEDLEKVLTKNQYQTERVFFYRAIASNTN 1086
 DB 779 FPY---EIRDEEAYLLEBETLENDYPLAWNLYLMEFKEALEKRSLOGRNPKWYQGRSQSL 835
 QY 1087 ER-----TWISTLSPGNCVNSIYNDEKTPISLYKKLFIISFNSFVDFLLRR 1137
 DB 836 SFPHDKKELIWTATKPPYVLDNRNLLFTGGNGP---YYGLINOSIYSLHYF----- 886
 QY 1138 FVDSNVKSLCYQCPMPQPEEKEILSNPLYLNLAKNTSLLIANKNDPENFKYLLYLYBFKF 1197
 DB 887 -----LGILSHPVIESMVKARA-----SEFRG-----SYSH 913
 QY 1198 DKEKNKI---LKLDEKDEPFK-----EK-----ENENN----- 1223
 DB 914 GKQFIEKPIRKIDFDQDQDEVKYNVTVTVEKLIITTDRIKESNGPRRRMLRRLDAL 973
 QY 1224 -----PIIASLYSLAKEDFITLLGD 1243
 DB 974 SNQLIQVINELYNISDEEYTVTLND 998
 RESULT 11
 ADU24023 standard; protein; 1009 AA.
 XX ADU24023;
 XX AC AC
 XX DT 27-JAN-2005 (first entry)
 XX DE *Bacillus pumilus* BpmI restriction endonuclease protein, BpmIRM.
 XX KW Type IIG restriction endonuclease; molecular cloning;
 XX KW gene characterisation; BpmI restriction endonuclease; BpmIRM; enzyme.
 XX OS *Bacillus pumilus*.
 XX PN US2004219584-A1.
 XX PD 04-NOV-2004.
 XX PF 15-MAR-2004; 2004US-00800946.
 XX PR 20-OCT-2000; 2000US-00693146.
 XX PR 17-MAY-2002; 2002US-00150028.
 XX PA (XUSS/) XU S.
 XX PA (KOB) KOBBE D.
 XX PA (ZHUZ/) ZHU Z.
 XX PA (SAMU/) SAMUELSON J.

PD 04-NOV-2004.
 XX 15-MAR-2004; 2004US-00800946.
 XX 20-OCT-2000; 2000US-00693146.
 XX 17-MAY-2002; 2002US-00150028.
 XX (XUSS/) XU S.
 XX (KOBBI/) KOBBE D.
 XX (ZHUZ/) ZHU Z.
 XX (SAMU/) SAMUELSON J.
 XX Xu S, Kobbbe D, Zhu Z, Samuelson J;
 XX WPI; 2004-783790/77.
 XX
 PT Altering specificity of type IIG restriction endonuclease by transforming
 PT cell with DNA encoding catalytic domain and methylase N-terminal of first
 PT endonuclease ligated to methylase C-terminal and specificity domain of
 PT second endonuclease.
 XX
 PS Example 5; Page; 34pp; English.
 XX
 CC The present invention relates to a method of altering the cleavage
 CC specificity of a Type IIG restriction endonuclease (E) having cleavage,
 CC methylase (M) and specificity domains. The method involves transforming
 CC cell with DNA encoding catalytic domain and methylase N-terminal of first
 CC endonuclease ligated to methylase C-terminal and specificity domain of
 CC second endonuclease. The method is useful for cleaving DNA molecules into
 CC small fragments for molecular cloning and gene characterization and for
 CC overproducing restriction endonucleases by genetic manipulation. The
 CC present sequence is the Bacillus pumilus BpmI restriction endonuclease
 CC mutant protein (B88A). Note: This sequence is not shown in the
 CC specification but is derived from Bacillus pumilus BpmI restriction
 CC endonuclease wild type protein shown as SEQ ID NO:4 in page 16-18 of the
 CC specification.
 XX
 SQ Sequence 1009 AA;
 Query Match 5.8%; Score 376; DB 8; Length 1009;
 Best Local Similarity 19.3%; Pred. No. 3.6e-13;
 Matches 260; Conservative 199; Mismatches 418; Indels 468; Gaps 55;
 14 ISLYKQLSKDFIKK-NINKLKDFFI--LIKNKLFSDIDNSTANIESLKYIFELNYS 70
 7 VDKYKAHRSTFLKPTNETQLRDFIDPLKSLGWDVDTKGTTHI---LRDVIQEEYIE 63
 71 VEQQKAGQIGVESRDVILLFENDKDKASFNKLEKAKKNEPIPIEDILIIAEVKRPTF 130
 64 IKD-----EETKKNPDYTLRINGTRK-----LFVAVKKPSF 94
 131 SPDAKDKLKESED--QLYRYLNOYQKHVGLSKVWRLYDKSKVLYGEXRYTEFNFXKI 188
 95 NI-----LKSAAKAFQRRRGWGANLGSIVTFHEHLVIYD-----CRY-----T 134
 189 EEKEEYKEQEWFLFYILIRKRYLKTNSVIEKEQIAKEBIQKTLKEILYERPDDS 248
 135 PDKSDNEHARYKVFSVEEYEAFADEIKDIISYESANSALDEMFVDNTR--VGETFDEY 192
 249 IVFKIAKIVYDKRFGSKGKITQHLIASILEESIIFIL-----RIFPIAVIENDNFKKI 303
 193 FLOQI-ENWREKLAKTAKNTE-----LGEEDVNFIVQRLNRIIFLRVCEDRTEIKY- 245
 304 LQENKLYRSSISFRYFFYDENTKKLRYKIIITENLLDKGSDAIKFPVNGGLFSEDKV 363
 246 -----ETIKSKNYEELKOLFQKSDR-----KFNGLFD----- 274
 364 KYNLNEGGLISSEIEELIVRM--LFPPEKNIKDEKVKYSLRDPKSGFELYETLLLEYDL 420
 275 -FIDDTLLLEVEDSNVLIIFPSDLYFPQP-----YDFSVDPTLSQIYERFLQEI 327
 421 RIADTTVHRHIEDGVYLIRTEEBELEKNVKNKIATYLGNIYLSRSRLDKKSGAYTTPDD 480

Db 328 -----IIESG-----GTFHITESPEVAASNGVPTPKI 355
 QY 481 LTDFMV---ISSIEQLTKSPDLIKIIDNSCGSHFLISCLDYLFYKFWYELDKFEDVK 537
 Db 356 IVEQIVKDTLTPTEGKFNELCNLKIADICCGSGTFLISSYDFLVEKV-----ME 406
 QY 538 KELDEEYRVIIIESE-EYDVQDSISKEIVLKMMLKXCIYGVINDIVSIEITWLSLWINT 596
 Db 407 KIIKEN-----IDSDLVYETEGELILTLKAKRNILENNLFGVDVNPYAVEAEFSL--- 458
 QY 597 FIFGTPLSFIEHRIKAGNALLGYTDEFFDIVKKPESGSLFKKRIKIIITILEDIYQK 656
 Db 459 -----LKLLE----- 464
 QY 657 IKGINDTTKEDIEKKIYKYEESKDIDNLRIRIIFSLIKLYSLSPKSNLMEPSDASVI 716
 Db 465 --GENEASVNN-----FIIHEHE-----DKIL-----PDLTSII 490
 QY 717 SLIENILGNK-----TSSDEKIEKIRKLSYKFFHYGIEFPDIOE--GFDIVGNP 768
 Db 491 KCGNSLVDNKFFEFMPESLEDEILFKANP-----FEWEEFDPDIMANGGFDIIGNP 543
 QY 769 PWEKTFNETEFSSKHIPNYRKLGIKEQNIKQEILSKDNHPLSIEYNEEKNIIAINNI 828
 Db 544 PY-----VRIQNM-----KYSPEEIEYYQSKDE----- 568
 QY 829 YKPDFKCFKFTGGDPLNRYFYFTFN--LKLIEKGNLTLYLPSAIMNESSRILRKHIFAR 886
 Db 569 -----YTVAKETVDKVFIERALILNLTGGLGYIIPHKFFITKGKELRKFAEK 621
 QY 887 YKLYNYIQENKRRKFDVHSSFKFAIFOLSNIKESTSSFKAKFMIOSSDNILKEITDLK 946
 Db 622 HOISKIINFGVTQVFPFRATYTAILIQANKM-----AQPKYKKNISAEITL--- 669
 QY 947 DSKDD--AYKGLTELNIQIKLSPIQESIIIEFKDNEETLINMKFSGALSGYIDFKK 1004
 Db 670 DSEENTCVISSEKYNSDPMIFLSPETEAV-----FTKFTB--AQFEKLGE-ITDISV 718
 QY 1005 GL-----DPSIKNRKSLKE-----C-----NNKNLI 1026
 Db 719 GLQTSADKIYIFIPENETSDTYIFNYKGRYEIEKSCCPAIVDLSFGSPESIQGNAQMI 778
 QY 1027 FLYSGANIIHQFSRPFEDKAKESKLLWIDKEDLEKVTKONQYOTERFVYRAIASNTN 1086
 Db 779 FPY---EIRDEEAYLLEETENDVPLAWNLYNEFEKLEKSLQGRNPKWYQYGRSOSL 835
 QY 1087 ER-----TWISTLSPGNCYCNSIYINDEKTPISLYKKLFIISFNSFVDFLLR 1137
 Db 836 SRPHDKELIWTVLATKPPYVILDRNNLLFTGGNGP---YYGLINOSIYSLHYF----- 886
 QY 1138 FVDSNVLKSCLYQCPMPQPEEKEILSNPLYLNLAKNTSLIAKNNDPENFKYLLYBYFKF 1197
 Db 887 -----LGILSHPVIESNVKARA-----SEFRG-----SYSH 913
 QY 1198 DKEKNVKI---LKDKEDEFFK-----EK-----ENENN----- 1223
 Db 914 GKQFIEKIPIRKIDFDQDEVKYNVTWTVTEKLIITTDRIKSESNGPRRRMLRRLLDAL 973
 QY 1224 -----FIASLYSLAKEDFITLLGD 1243
 Db 974 SNQLIQVINELYNISDEEYTTVLND 998
 RESULT 13
 ADC00979
 ID ADC00979 standard; protein; 1644 AA.
 XX
 XX ADC00979;
 XX AC
 XX DT 04-DEC-2003 (first entry)
 XX DE Enterohaemorrhagic E. coli 0157:H7-specific protein SEQ ID NO: 1023.
 XX

KW enterohaemorrhagic; anti-bacterial.
 XX Escherichia coli; 0157:H7.
 XX JP2002355074-A.
 XX 10-DEC-2002.
 XX 24-JAN-2002; 2002JP-00015959.
 XX 24-JAN-2001; 2001JP-00112010.
 XX (UYTS-) UNIV TSUKUBA.
 XX WPI; 2003-451640/43.
 XX Enterohemorrhagic Escherichia coli O157:H7-specific nucleic acid molecule
 PT and a polypeptide and its use, a polypeptide, a vector and a host cell.
 XX Claim 3; SEQ ID NO 1023; 2067pp; Japanese.
 XX The invention relates to a novel enterohaemorrhagic Escherichia coli
 CC O157:H7-specific nucleic acid molecule. A polynucleotide of the invention
 CC has anti-bacterial activity. The polypeptide can be used in detection
 CC and/or treatment of O157:H7 infection. The nucleotide sequence of the
 CC genome of Enterohaemorrhagic E coli O157:H7 was determined. The present
 CC sequence represents an E. coli O157:H7-specific polypeptide of the
 CC invention.
 XX SQ Sequence 1644 AA;

Query Match 5.7%; Score 371.5; DB 7; Length 1644;
 Best Local Similarity 17.7%; Pred. No. 1.2e-12;
 Matches 246; Conservative 231; Mismatches 470; Indels 443; Gaps 51;

QY 74 QKAGQIEGVSRRVILLFENDKDKAFNNKLEAKKNEPIEDIL---IIAEVKRPTF 130
 DB 143 QEEGTLDPLALSLLTAQFPADTDKHDLSLRKA--NGEYRSWQDLLSTAVFTQNEPPRF 200
 QY 131 SFDANKLKESEDQLYRLNQYKHGILNGKWRLYDKSVLYGKRYIEFNFKIEE 190
 DB 201 V-----LLGNRLQLLLDRTK--WAQNRLLRFPDFEILS 232
 QY 191 KEEYKEQEWFLVILIRKERYLKTNSVIEVEKEQIAKEKEI--IQKTLKEILYERPDSS 248
 DB 233 R---REYDTTLKATAVLLHKDSLPGSGAPYLDSDLDNSHSHKAFGVSEDLTALRE----- 284
 QY 249 IVFKIAKNIYDKFKLGGKEITQHIL-----ASTLEESIIFILRIFFIAY 293
 DB 285 -----SIELAGNEAMHYLIDRGLANYTGNAVDPDDELRECLAYMYRLFLFY 332
 QY 294 IENDDIPFKILOENKLYRSSISPRY-----FFYDENTKKLEYKKIITIFNLLDKGS 345
 DB 333 IEARPELGAPYMAKTYLQGYSLFTRDLEMLPITSEEDRNGRYFHDLSNMLFKLVREG- 391
 QY 346 DAIKFPVPNGG-----LFSQDKVLYNLEGLLSISEIE-----EILVKMLRPEEK 390
 DB 392 -----YNGGVKMQSLESQDRITIIHSQ--FSVPRLESHLPDANNTRILNRVFRNET 442
 QY 391 -----NIKOEKFKYKSLRDPKSGELYETILLEYDLRIADTTHRI----- 430
 DB 443 LQOIIQAMSLSRPAKGFNRGRISYRQLGINQLGAVYEALLSYRGFFASEDLVEYKKG 502
 QY 431 -----IDGVVLIHT-----EELENKKNKIATYKGNLYLTSRLDRKKGAYYTPD 479
 DB 503 EEFNELETGVFVSKDEISKYHEDSKYKDGSLRIHRKGSFIYRMAGDRKESASYYTPE 562
 QY 480 DLTDFWVISSTEE-----QLKTKSPLDIKIDNSCGSHFLISCLDVLTEKVV 527
 DB 563 VLTRSLVKYALKELFKGQIDIPSHAKADAILMLTVCEPAMGSAALFNAIQLABAYL 622
 QY 528 YELDKFEDVKKELDEEYRVIIIESEEDVDQDSISKELV-LKRMLLKXCIYGVDPINPISVE 586

DB 623 FH-----KQAEGRRI-----PDQRYTQELQVKMYIADNNVFGVDLNPVAVE 665
 QY 587 ITMLSLWINTFIQTPLSFIEHHIKAGNALLG-----YTKDEFF----- 625
 DB 666 LAEVSMLNLAISGDAFVFWPGYQLHCGNSLVGARQVFNKSELTYYKAKDPWLNSEPYE 725
 QY 626 -----DIVKKKPSGFSLSFKKRIKEI-----ITIL 650
 DB 726 LAMNTPRETOIFHFLPDGGMANYSDKTVKQRYDDPKALDSWRKEFTKSFAGHEIADV 785
 QY 651 EDIYQIKIGINDTTKEDIE-----KSKKIYKEYESKD 683
 DB 786 QRISEKVEALMNTYRQLKAERLKTADSYVPMVPAENSEQTRSLSKDETFSGRLDNSA 845
 QY 684 IDNLRITFSL-IKLYSLSPDKSLNMEFSDIASVISLIENILGNKTSSEDEKIEK----- 737
 DB 846 YQKLKRWMDYWCALMFWPIDKA--DELPRGTWLFIEITLLDGLVITEKTEVAEHTTGD 903
 QY 738 -----IRKLSYV-----KPHYGIEPPDI---QE 759
 DB 904 LFAEGLLRRESSLFSVAGRLKTEVLFRHLPRLAIVDALRKHQRPFFHMDLEFCDLFAERG 963
 QY 760 GFDVIGNPNWETKFNTEFFSKHPIYV--RKLGIKEQNIIOEILSKDNHPLSEYNE 817
 DB 964 GFDMLGNPPWLKVEWQEAAGVLDYEPEFVLRKLSASKLATLRIDTF----- 1010
 QY 818 EKNSIIAINNIYKDFK-C-----FTSGGDPNLFYVPTVFNKLKIEKGNL 862
 DB 1011 --NQIPALEAAWRSEYEGCEGMQNFNAQONTVPLRGVQTNLYKCELPQAWRLGAQKQVA 1068
 QY 863 TYLVPSAIWNSESSRIILKHIIFARYKLIYIOFENKRRF---KDVHSSPKFAIF----- 913
 DB 1069 GFHLPEGIYDDPKGGQLRAAVYPRLAH--FQFQNELNLFVEVDHHAKESSNIYASPSPT 1126
 QY 914 ----QLSNI-----KESTSSFKAKFMIOQS-----DNILKEITRD 944
 DB 1127 VGFEHISNLYAPQTTIDACFEHSGSGDIPGLKDEIESEGLKVVMWNTSGHRSRLSIATHE 1186
 QY 945 LK-----DSK-DDAYKG-----LELNINQIKLSPIQESIIIEFKDNEFTL 984
 DB 1187 LELFARLYDSEGTPAQARLPALHAKQLVALEKFAQNPRLGDLO-----GQVFST 1238
 QY 985 INKMFSGKSGALGEGYI-----DFKGLDPSIKNRKSLKCEKNKNLI 1026
 DB 1239 V--MFEETVQAQDGTILROTQFPDSSQWVLSGPHFVG--TPYKTPRENTLNSVDCL 1295
 QY 1027 FLYSGANIHFQNSRPFEDKAKESK---LWID--KEBLEKVLTKDNQYQTERVYRAI 1081
 DB 1296 DLLTLDPDYLPRNTYIPACDAQEYAKRTPCVTWTELAEDPEPKVT--DYV---RLAIRAM 1350
 QY 1082 ASNTNERTWISTLSPG-----NCYCVNSIYNDEKTPISLYKGLFIISFNSFVDF 1133
 DB 1351 LAQSGERTLISAIYPPEISHMNAVRSYSSQWMLLEHSGMCF-----SLPFD 1399
 QY 1134 LLRRFVDSNV 1143
 DB 1400 ICKSTCKANL 1409

RESULT 14
 ADN46345
 ID ADN46345 standard; protein; 1282 AA.
 XX AC ADN46345;
 XX DT 01-JUL-2004 (first entry)
 XX DE Thermococcus kodakaraensis KOD1 protein sequence SeqID223.
 KW gene disruption; gene targeting; marker gene; transformation;
 KW homologous recombination; hyperthermostable archaeobacterium; KOD1;
 KW gene structure; gene function; enzyme activity; medicine;
 KW forensic science; food; drug inspection; molecular biology; immunology.

XX OS Thermococcus kodakaraensis.
 XX PN WO2004022736-A1.
 XX PD 18-MAR-2004.
 XX PF 29-AUG-2003; 2003WO-IB003597.
 XX PR 30-AUG-2002; 2002JP-00319011.
 XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 XX PI Imanaka T, Atomi H;
 XX DR WPI; 2004-257583/24.
 XX PT Method for disrupting targeted gene in genome of organism particularly
 PT thermostable bacterium and with genome chips for analysis, applicable in
 PT studying gene structure and functions.
 XX PS Claim 9; SEQ ID NO 223; 598bp; Japanese.
 XX CC This invention relates to a novel method for targeting disruption of an
 CC arbitrary gene in a genome of an organism which comprises providing the
 CC whole sequential data of the genome of such organism, selecting at least
 CC 1 arbitrary region in the sequence, providing a vector that contains a
 CC sequence homologous with the selected region and a marker gene,
 CC transformation, and homologous recombination. The genome is preferably
 CC the genome of a hyperthermostable archaeobacterium, particularly
 CC Thermococcus kodakaraensis KOD1. The method is for targeting the
 CC disruption of a gene in the genome of an organism, which is applicable in
 CC studying gene structure and functions as well as enzyme activities of
 CC encoded proteins and useful in medicine, forensic science, food or drug
 CC inspection, molecular biology and immunology. With this method, the
 CC disruption of a gene at an arbitrary position in a genome can be achieved
 CC efficiently and reliably. The present sequence is that of a protein
 CC encoded by the genome of Thermococcus kodakaraensis which was derived
 CC using the method of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 1282 AA;
 Query Match 5.6%; Score 365.5; DB 8; Length 1282;
 Best Local Similarity 18.9%; Pred. No. 1.9e-12;
 Matches 265; Conservative 245; Mismatches 413; Indels 481; Gaps 68;
 QY 121 IIAEVRKPTSPD-AKDKL-----KESDQLRYRLNQYQKHYGILSNQWRLYDKSK 172
 DB 96 IIVEFKR---LPFLRRNKLARTDLNWKHEEQIKYVFSRFPVLTNLVYVFFSKT 152
 QY 173 VLYGKRYIEFNPAKIEKEEYKQEWFLVPIVLRK-ERYLTSNVIEVEKEQIAKKB- 230
 DB 153 I-----IRKSRFP---YYATFDLVHDIEQYGSGLYDILLERKEHGIQKGL 194
 QY 231 -ELIQTKLELYERPDSDIVFKIKNIVDFEKLKSGKEITQHILASILEESIFILRIF 289
 DB 195 DNRFFKSLREWI-----KTLEDVEFK-TDERIKMHLHLINK-FIFIOTLD 239
 QY 290 FIAYIENDIDFKILQENKLYRSISPRYFFYDENTKKKLE-----YKKIITIFNLIDKG 344
 DB 240 DYGVVERNLVKKWGMGFSRILRA-----ERHKEPEVAGRHYKSFIEKF--LGEI 288
 QY 345 SDAIKFPVFNGLFSEDKVLYLNN-----EGLLSISEIEILVLMJFFKEKNI 392
 DB 289 NDFF-YPLVTELFSDRLIEFIKDEPENWARFYKALETVLGFAPWQETGLKMGITQ---- 343
 QY 393 KDEKFKYSLDPKSPCELYETLLELDLRTADTVVHRIIEDGYLIRTEELEENKVNKI 452
 DB 344 -----YDYSQIDEDILGRAYET----- 360

QY 453 ATYLKGNIVLTSRLDRKSGAYYTPDDLTDFTWVSSIEQLK----- 495
 DB 361 -----YLABK---RKEGIIYTPKYVTQFIVEETLQKLNELKEGITTARENDPFT 409
 QY 496 ---TKSPLDIKIIDNSCGSHFLISCLDYLTKEVWYELDKPDKVKEL-----DEEVRVI 547
 DB 410 AERLSKELEIKVADLASGSGFLIKVLRGLWDTYSAVIDELKKKEELIGKTSLSAL 469
 QY 548 IESEEDYVQDSISK-----ELVLRMLLKCIYGVNDINPISVEITMISLWINTFTI 598
 DB 470 VQRK---DIIEGIAKMRLEFPDSERILMSQVLRH-VFAVDLDENAVAKNMLW-RELI 524
 QY 599 FGTPLSP-----IEH-----HIKAGNALLGYTKDEFF---DIVKKKFSGFSLSFK 641
 DB 525 KLNPKAFRWDNLGENEHLPLNLVSGDLSGLFADPKILEGREEVRLME---LWEK 580
 QY 642 RIKEIITILEDIVQIKIGINDTTKEDIKSKKIYKEYEESKODIDNLRIRIIFSLIKYLSLF 701
 DB 581 FIENPEDL--EVLNEIKEIKETLRGELDER---YRE----- 611
 QY 702 DKSLNMEFSDIASVISLIENILGNTSSSEDKIEKIRKLSYKFFHYGIEF-----P 755
 DB 612 -----LLREKLG-----EKAELANN---RFIHPLEFMAFFNP 642
 QY 756 D--IQECPDIVIGNPPW-----EKTFPNE--TEFFSKHIPNYRKLGI-----KEQNII 799
 DB 643 DGLSRGGPFIIVGNPEYVRIQNLKKESEPEYFLNRYSAHKNY-DLAIPFTIERGYNLL 701
 QY 800 KGE-----ILSKDNHPLSEYNEEKNISIAINNIYK--PDFKCTSGDGNLPFRYFTFN 852
 DB 702 REGGELGFIVTK--KWKIDYGEKLRGLSRERAVRLIIDF-----GDEQVFKGATTYT 753
 QY 853 LKLI--KEKGN-LTYLVPSAI-----WNSSSRIL-----RKHIPA 885
 DB 754 MILVLRKERNDKLTIAKVEELKETVDQLRAVNPDKWNSERLSVIKPTKELSEKWPVL 813
 QY 886 RYKLVYIYO--FENKRFKDVHSSP-----KFAIPOLSNIKESTSFPKAKMIQSSDNI 937
 DB 814 TEEKKIVKIKYNGSKRLGEVTDIFVGLQTSADPVVILEYVSGESGNYIIVHVKITQOEHR 873
 QY 938 LKEITRDLKSDKDDAVKIE----- 957
 DB 874 LE-----KOLLHPLLKGEIRRWMIPEYRYLILPPYKVLVNGERKABELIPIEELASKY 927
 QY 958 -----LNINQIKKL-----SPIQESIIEFKONEEF---TLINKMFS---KFS 994
 DB 928 PRIVEYLNISEVRKLEGRGSLKSGSPKWSYVVEKNEHEKPELPKLTITGVLSSEPRFAL 987
 QY 995 LGEYIDFKKG-----LDPSIKNRKSL--LKECNRNLI-----FL 1028
 DB 988 DKNGDYVYFVGSGTAGGYGVLIKDDYKDRISLOFIALNSSLDMRWVKQITGSEFEGFYS 1047
 QY 1029 YSGANIHQNSRF---FEDKDAKESKLWIDKEDKVLTKDNQYQTERVFRATASN- 1084
 DB 1048 YGRASIKOLPIKLPTDEESTLKBEI-----EDTVEEIIALLKHHVYVKSRLREWSKKL 1101
 QY 1085 TNERWTISLSPGNCVNSIYINDEKTPISLYKKLFIISI-----FNSGFVD 1132
 DB 1102 ANKGTTLGLIEGWKKGUG-----KLP---PEKLTFTDRLISNEEMEYEGFELEKLD 1151
 QY 1133 FLIRRF-----VDSNVLKSLCYQCPMPQPEKEITLSNPLYNLAKNTSLLI 1178
 DB 1152 GTLRLLFGREWDMLPTILELEGDEEILEH-VYLSMLELLESRRQVKN--LKDILNKTIRVLT 1208
 QY 1179 AKNDPENPKLYLYEYFKPDKKVNKILKLDKDEDEFFKEKENENNFIIASLYSLAKEDFI 1238
 DB 1209 IGNDEPETERIRIVKTVQRAGVRLTSLFLELVRENEAYLDA-----LVFKLYGLSTEER 1262
 QY 1239 TLIGDGFALKNNKKGDEYISSLIK 1262
 DB 1263 IV-----LRNLNKTQDYIDSVR 1280

RESULT 15
AAU48711
AC AAU48711 standard; protein; 1629 AA.
XX
DT AAU48711;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #9607.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypotosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
XX WO200181581-A2.
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US012865.
XX
XX 21-APR-2000; 2000US-0199047P.
XX
XX 02-JUN-2000; 2000US-0208841P.
XX
XX 07-JUL-2000; 2000US-0216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
XX
XX N-PSDB; AAS59543.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
XX
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris.
XX
XX Example 1; SEQ ID NO 9906; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX polypeptides. The proteins and their associated DNA sequences are used in
XX the treatment, prevention and diagnosis of medical conditions caused by
XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX pustulosis, hypotosis and osteomyelitis), uveitis and endophthalmitis.
XX P. acnes is also involved in infections of bone, joints and the central
XX nervous system, however it is particularly involved in the inflammatory
XX lesions associated with acne vulgaris. A method for detecting the
XX presence or absence of P. acnes in a patient comprises contacting a
XX sample with a binding agent that binds to the proteins of the invention
XX and determining the amount of bound protein in the sample. The
XX polypeptides may be used as antigens in the production of antibodies
XX specific for P. acnes proteins. These antibodies can be used to
XX downregulate expression and activity of P. acnes polypeptides and
XX therefore treat P. acnes infections. The antibodies may also be used as
XX diagnostic agents for determining P. acnes presence, for example, by
XX enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1629 AA;
Query Match 5.6%; Score 364; DB 4; Length 1629;
Best Local Similarity 16.5%; Pred. No. 3.2e-12;
Matches 242; Conservative 246; Mismatches 531; Indels 446; Gaps 52;
Qy 70 SVEQQKAGQIEGVESRDVILLFENDKOKASFNWKLKAKK--NNEPIP-----IEDILI 121
Db 86 SVPMPSDDALIVVDWISSEHFTTDDARKESQKLVLDRKQWQDGEVPTTTRPTERASK 145
Qy 122 IAEVKRPTFSFADKDLK-----ESEDQLRYLQYQKHGILSNGKV----- 164

Db 146 LATMLAALYSEDLDAAARAAEAQAHELLRVLGSTGFEIRRDGPFVWFSTAGIEQSA 205
Qy 165 -----LYDKSKV-----LYGEKRYIEF-- 183
Db 206 LAVVTARPVETPDQLMVKDARTLPTFPRSEDPADSBSEVHSTSRILVSTLIFAPKEGSPFAL 265
Qy 184 ---NFXKIEEKEVEKEQEWFLVYLIRKERYLKTSNVIE-----VEKEQIAKE----- 229
Db 266 IMAGRWLVVAEGRWPEGRYLAVDQTVARADLTKGGEVDRLATCLEASLAPDADGHV 325
Qy 230 --KEIIQKTLK-----EILYERPDSDIVFKIAKNIYDKFELKSGKEITQHILASILEBS 281
Db 326 WNSEALDESIKHTVGVSDQLREGVRESIEI-IANEVVSRRRDRGLEPLSDQDOQAPLAMOS 384
Qy 282 IIFILRIFFIAYIEDN-----DIFKKILQENKLYRSSISFRFFYFD 322
Db 385 LRFLYRILFLYAEASPELGVLPVGAPOVDSYSLDRRLRELQVPLVLTQRSMAAGTHL--- 441
Qy 323 ENTKKKLEYKIIITIFNLDDKGS-----DA---IKFPVFNGGLFSBDKVKYLNNEGGLSI 374
Db 442 -----YESIGTLFRLVDRGNLEDDVDTTGLSFHSLRADLFRQATALIDEVG-LGN 492
Qy 375 SEIEEILVKOLFEE---KNIDEKFKVYSRLDPKSFGELEYETLLEYDLRIADTTVHRII 431
Db 493 QALQVLERLLLSKEQHGRKRGYISYAEGLINQLGAVYEGLSYSGFFATEDLYEVA 552
Qy 432 EDG-----VYLIRTEEELENKVNKI-----ATYLGKNIYLTSLDRKKGAY 475
Db 553 HDGNPQKGSWVPLSRAHDIAQEDFTVVVDHGHARPHIRHRRGEFVRLSGRARQOSASY 612
Qy 476 YTPDDLTDFWJSSIEEQL---KTKSPLDIKIINSCGSHFLISCLDYLTKEKWWELD 531
Db 613 YTPVLTFRVGOALEELDDTTTAEIHLHLSVCEPALGSGAFATQALQAEQ--YLTR 670
Qy 532 KFEDVKKELD-BEYRVIIIESEYDVQDSISKELVLRMLLKXCIYGVDPINPISVITWML 590
Db 671 RQAEIGRRIDPDYPALQORTRAY-----IALHNVEVDNLNATAVEFAEI 715
Qy 591 SLWINTFFIGTPLSFIEHHIKAGNALLG-----YTKDE-----FFDIVKK 630
Db 716 TLWLDTMASGLDAPWFLHRRGNSLIGARHAYVTRDQLTSKAWLTPTTPTDVLTLTLAAR 775
Qy 631 KRES-----CFSLFKKIKIETIILEDIYOKIKIGINDTKE--- 666
Db 776 LDDDAVDPTQAGSIHPLVPAAGWSSADSKEATLVP-----QRVKDKWRKQFRR 829
Qy 667 ---DIEKSKIKIYEYEE-----SKDI----- 684
Db 830 RPSRHVDVTLVELGHQVEDLWTMALRRLTVAEQTSRDIALMGRNEAPSHRQVVTREQE 889
Qy 685 -----DNRIITFSLIKLYSLSPDKSLNMEFSDI-----ASVISLIENILGN 725
Db 890 QSLADPDGAYQRLRLV-----MDAWCALWFMPPLTTQVAPPSPWAEWIDACTMLLGG 939
Qy 726 KTSSEDKK-----ISKIRKLSYYK 746
Db 940 RLBRADRRDAPMLGSAQIWEALAAQEDLALASAGATVAQTLDKHPWLOVCRQVAREQ 999
Qy 747 FFHYGIEFFDI--QEGFDIVIGNPPWEKTKFNETFFSKHIPNYRKLGIKEQNIHQEIL 804
Db 1000 FFHWQLDFAPVFARGGDFLQVGNPPVRSILDMALLAEGDPWQLAGKTPEDVREQ--- 1056
Qy 805 SKDNHPLSTIYNEEKNSITAINNIYKDFKC-----FTSGGDPNL 844
Db 1057 -----RRTVTTLAPGL--ADSVCRGITDVGSGSFVGHATNPMQLGQPLD 1101
Qy 845 FRYFVTFLNLKLIKEKGNLTVLPVSATWNSSRIIRKHIFARYKLNVIYOFENK-KRFXD 903
Db 1102 YRCFMQOTWHTSSRGITGLVHPETHFTDEKAGHLREETVPLRRH--WQFVNLKLFDE 1159
Qy 904 VHSFSPKFAIPQLSNIKESTSSFKAKPMIQSSDNILKEITRD----LKSDKDDAYKGIELN 959

Db 1160 VHDLVTVGVHUYGS--PAQPHFLOASALYHPDPTVVGSLRHDGSGGAPGPKVDGHWDRPH 1217
Qy 960 INQIKLSPIQESIIEFKD-----NEEFTLINKMFSKFSALGEGYID 1001
Db 1218 AQRLETVT--DETLATWRDILDPNLEAPRRTREMLYTVNRDVAETLQOLSKAPRLGSLSLR 1275
Qy 1002 FKKGLDPSIKNRKSLKE-----CNNKXLIPLYSGANIHQNSRPFEDKDAKESKLLW-- 1055
Db 1276 FSPGWHEKNORTKGYFIQQWGTDPDSWNDVIL-QGPHLH-VATPFYKSPNPTMKHNDQWSV 1333
Qy 1056 IDKEDLEKVLTKONQYOTE-----RVFYRAIASNTNERTMI-- 1091
Db 1334 VDLETLPDAIPVTSYKPGADRAWYDADYTHWDGDPARDHYRLAWRAMAANTGERTLIPA 1393
Qy 1092 ----STLSPGNCYCVNSIYINDEKTPISLYKKLFIISIFNSFVDFLLRRFVDSNVLKSC 1147
Db 1394 IIPGTAHPNGVFCVGA--DNRILTA-----CAGFASLLLDPSARAAPKSGIYOAV 1444
Qy 1148 LYQCPMP---QPEKEILSNPLYLN 1169
Db 1445 FDRLPAPCORHPLLPALLRLRLN 1469

Search completed: January 24, 2006, 19:44:49
Job time : 299.376 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:28:54 ; Search time 48.4458 Seconds
(without alignments)
2536.214 Million cell updates/sec

Title: US-10-688-058-16
Perfect score: 6538
Sequence: 1 MKTNDIVKTNPNISLYKQL.....SSLIKGYNLLNNKIFVHK 1277

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: Pir1:*
2: Pir2:*
3: Pir3:*
4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6534	99.9	1277	2 E70224	hypothetical prote
2	5599	85.6	1278	2 B70236	hypothetical prote
3	1781	27.2	1250	2 E81339	probable restricti
4	1648.5	25.2	422	2 D70251	hypothetical prote
5	558.5	8.5	706	2 T08313	conserved hypothet
6	526.5	8.1	1181	2 B64516	hypothetical prote
7	473	7.2	134	2 E70251	hypothetical prote
8	455.5	7.0	1252	2 D71810	probable type II D
9	446	6.8	867	2 B81419	probable type IIS
10	419.5	6.4	1279	2 E64709	type IIS restricti
11	378.5	5.8	2401	2 T28676	rhostry protein -
12	371.5	5.7	1644	2 F91286	hypothetical prote
13	371.5	5.7	1644	2 B86128	hypothetical prote
14	361.5	5.5	1339	2 H81307	restriction modifi
15	359	5.5	2166	2 G70163	hypothetical prote
16	357	5.5	2269	2 T28677	rhostry protein -
17	349.5	5.3	1979	2 C71622	hypothetical prote
18	342	5.2	4981	2 T18489	hypothetical prote
19	336.5	5.1	2819	2 A90551	conserved hypothet
20	336.5	5.1	3394	2 T18501	hypothetical prote
21	336.5	5.1	3973	2 B71612	hypothetical prote
22	335.5	5.1	1939	2 T18372	repeat organellar
23	333.5	5.1	993	2 A97334	type II restrictio
24	333.5	5.1	1127	2 T28317	ORF MSV156 hypothe
25	330	5.0	2136	2 A05037	hypothetical prote
26	328.5	5.0	3724	2 T18427	hypothetical prote
27	326.5	5.0	2295	2 B71621	probable membrane
28	320.5	4.9	1650	2 T18444	hypothetical prote
29	316	4.8	1639	2 S05603	major merozoite su

30	313.5	4.8	1183	2 F90559	conserved hypothet
31	312	4.8	1516	2 E71619	RAD2 endonuclease
32	308	4.7	1631	1 SAZQK1	major merozoite su
33	308	4.7	1640	2 A24594	probable major sur
34	307	4.7	2485	1 H71621	serine/threonine-s
35	306	4.7	2829	2 A42771	reticulocyte-bind
36	305.5	4.7	1787	2 G97222	hypothetical prote
37	305	4.7	3119	2 T18414	protein g377 - mal
38	302.5	4.6	1619	2 T18499	hypothetical prote
39	301.5	4.6	1163	2 G97236	ATPase involved in
40	301	4.6	1701	2 A26868	major merozoite su
41	298	4.6	1018	2 E64451	type I restriction
42	298	4.6	2013	2 C71610	probable membrane
43	297.5	4.6	609	2 D75633	conserved hypothet
44	297.5	4.6	971	2 A70179	exodeoxyribonuclea
45	295.5	4.5	1875	2 S38173	myosin-like protei

ALIGNMENTS

RESULT 1

E70224
hypothetical protein BBE02 - Lyme disease spirochete plasmid E/lp25
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C:Accession: E70224
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
i Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390: 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A>Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: E70224
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1277 <KLE>
A:Cross-references: UNIPROT:O50698; UNIPARC:UPI0000056833; GB:AE000785; NID:g2689951; PI
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid

Query Match	99.9%;	Score 6534;	DB 2;	Length 1277;
Best Local Similarity	100.0%;	Pred. No. 1.8e-241;		
Matches 1277;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MKTNDIVKTNPNISLYKQSKDFIKKENINKLKDFILIKNKLFSIDDNSTEANIESLL	60	
Db	1	MKTNDIVKTNPNISLYKQSKDFIKKENINKLKDFILIKNKLFSIDDNSTEANIESLL	60	
Qy	61	KYIFEELNYSVEQKAGQIEGVESRDVILLFENDKDKASFNNKLKEAKKNEPIPIEDIL	120	
Db	61	KYIFEELNYSVEQKAGQIEGVESRDVILLFENDKDKASFNNKLKEAKKNEPIPIEDIL	120	
Qy	121	IIAEVGRPTFSFADAKLKESEDQLYRYLNQYQKHGYLSNGKVRWLYDKSKVLYGKRY	180	
Db	121	IIAEVGRPTFSFADAKLKESEDQLYRYLNQYQKHGYLSNGKVRWLYDKSKVLYGKRY	180	
Qy	181	IEFNFXXKIEEKEEYKEQEFVLYLRKERYLKTNSVIEVEKEQIAKEKEIIQKTLKEI	240	
Db	181	IEFNFXXKIEEKEEYKEQEFVLYLRKERYLKTNSVIEVEKEQIAKEKEIIQKTLKEI	240	
Qy	241	LYERPDDSDIVFKIAKNYDKFKLSGKEITQHILASILEESSIIIFLIRFFIAYIEDNDIF	300	
Db	241	LYERPDDSDIVFKIAKNYDKFKLSGKEITQHILASILEESSIIIFLIRFFIAYIEDNDIF	300	
Qy	301	KKILQENKLYRSSISFRFFYDENTKKLEYKKIITIFNLLDKGSDAIFKPFVFNGLFSE	360	
Db	301	KKILQENKLYRSSISFRFFYDENTKKLEYKKIITIFNLLDKGSDAIFKPFVFNGLFSE	360	
Qy	361	DKVKYLNNEGLLSISEIEELVVKMLFPEEKNIKDEKFKVYSRLDPKSGELYETLLEYDL	420	

361 DKVKYLNNEGLLSISIEBIEILVMQLFFEEKNIKDEKFKVYSRLDPKSGELYETLLAYDL 420
 QY 421 RIADTTVHRI EDGVVLTIRTEBELENKVNKIATYLGKNIYLTSSLDKSGGAYTTPD 480
 Db 421 RIADTTVHRI EDGVVLTIRTEBELENKVNKIATYLGKNIYLTSSLDKSGGAYTTPD 480
 QY 481 LTFDFWISSIEBQLTKSPDLKIIDNSCGSGHFLISCLDYLTKVWYELDKFEDVKREL 540
 Db 481 LTFDFWISSIEBQLTKSPDLKIIDNSCGSGHFLISCLDYLTKVWYELDKFEDVKREL 540
 QY 541 DDEYRVIISESEYDVQDSISKELVLRMLLKXCIYGVNDINPISVEITWLSLWINTFI 600
 Db 541 DDEYRVIISESEYDVQDSISKELVLRMLLKXCIYGVNDINPISVEITWLSLWINTFI 600
 QY 601 TPLSFTEHHI KAGNALLGYTKDEFFDVVKKESFSGSLFKKRIKEIITILEDIYOKIKGI 660
 Db 601 TPLSFTEHHI KAGNALLGYTKDEFFDVVKKESFSGSLFKKRIKEIITILEDIYOKIKGI 660
 QY 661 NOTTKEDIEKSKKIYKEYESKDIIDLRIIFSLIKLYSLSPDKSLNMFSDIASVISLIE 720
 Db 661 NOTTKEDIEKSKKIYKEYESKDIIDLRIIFSLIKLYSLSPDKSLNMFSDIASVISLIE 720
 QY 721 NILGNKTSSEDEKIEKIRKLSYKFFHYGIEFPDIOBGFQDIVIGNPPWEKTKFNTEF 780
 Db 721 NILGNKTSSEDEKIEKIRKLSYKFFHYGIEFPDIOBGFQDIVIGNPPWEKTKFNTEF 780
 QY 781 FSKHPNYKLGIGKEQNIIOBILSKONHPLSIEYNEEKNSIIANNIYKPKFCFTSGG 840
 Db 781 FSKHPNYKLGIGKEQNIIOBILSKONHPLSIEYNEEKNSIIANNIYKPKFCFTSGG 840
 QY 841 DPNLFRYFTFNLKLTKEKGNLTYLPSAIWNSSSRILRKHI FARYKLNLYIYQENKQR 900
 Db 841 DPNLFRYFTFNLKLTKEKGNLTYLPSAIWNSSSRILRKHI FARYKLNLYIYQENKQR 900
 QY 901 FKDVHSSFKFAIFQLSNIKESTSSFKAKPMI QSSDNILKEITRDLKDSKODAYKGIELNI 960
 Db 901 FKDVHSSFKFAIFQLSNIKESTSSFKAKPMI QSSDNILKEITRDLKDSKODAYKGIELNI 960
 QY 961 NOIKKLSPTOESIEIEKONEBFTLKNMFSKSGSALGEGYIDFKGLDPSIKNRKSLKEC 1020
 Db 961 NOIKKLSPTOESIEIEKONEBFTLKNMFSKSGSALGEGYIDFKGLDPSIKNRKSLKEC 1020
 QY 1021 NNKQLIFLYSGANIHFQNSRPFEDKAKESSKLLWIDKEDLEKVTAKONQYOTERVFYRA 1080
 Db 1021 NNKQLIFLYSGANIHFQNSRPFEDKAKESSKLLWIDKEDLEKVTAKONQYOTERVFYRA 1080
 QY 1081 IASNTNERTMI STLSPGNCYCVNSIYINDEKTPISLYKKLFIISIPNSVPFDFLLRRFVD 1140
 Db 1081 IASNTNERTMI STLSPGNCYCVNSIYINDEKTPISLYKKLFIISIPNSVPFDFLLRRFVD 1140
 QY 1141 SNVLKSLCQCPMPQPEERKEIISNPLYINLAKNTSLIIAKNDPENPKYLLYLEYKFKDKE 1200
 Db 1141 SNVLKSLCQCPMPQPEERKEIISNPLYINLAKNTSLIIAKNDPENPKYLLYLEYKFKDKE 1200
 QY 1201 KVNKILKDKDEBFFKEKENNNFIISLYSLAKEDFTLLGDFAKLNKKGEDYISSL 1260
 Db 1201 KVNKILKDKDEBFFKEKENNNFIISLYSLAKEDFTLLGDFAKLNKKGEDYISSL 1260
 QY 1261 IKGYDNYLLNNKIFYHK 1277
 Db 1261 IKGYDNYLLNNKIFYHK 1277

RESULT 2
 B70236
 hypothetical protein BBH09 - Lyme disease spirochete plasmid H/1p28-3
 C:Species: Borrelia burgdorferi (lyme disease spirochete)
 C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
 C:Accession: B70236
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
 son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
 i Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
 Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.
 A>Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
 A:Reference number: A70100; MUID:98065943; PMID:9403685
 A:Accession: B70236
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1278 <KLE>
 A:Cross-references: UNIPROT:O50667; UNIPARC:UPI0000056814; GB:AE000784; NID:g2690041; P1
 A:Experimental source: strain B31
 C:Genetics:
 A:Genome: plasmid

Query Match 85.6%; Score 5599; DB 2; Length 1278;
 Best Local Similarity 85.3%; Pred. No. 6.6e-206;
 Matches 1082; Conservative 94; Mismatches 89; Indels 4; Gaps 3;
 QY 6 IVKTNPNISLYKQSLKDFIKKENINKOPFIIKKNLFSIDDNSTANIESLLKYIFE 65
 Db 12 IIKTNDPNVSLYKELSGFIKENIVKSNKFFILKKNKIQIADNSTANIESLLKSIFE 71
 QY 66 ELNYSVEQOKAGQIEGVESRVDILLFENDKDKASFNKKLKEAKKNNEPIEDILIIAEV 125
 Db 72 ELNYSVEQOKGGQIEGVKSRVDILLFENDKDKVAFNKKLBEAKKNNESIPTEDILIIAEV 131
 QY 126 KRPTFSFPAKDKLXESDQLYRYLNQYKHVIGILSNKVMWLYDKSVLYGKRYIEFNF 185
 Db 132 KRPSFSPAKDKVXAEADQLYRYLNQYKHVIGILSNKVMWLYDKSVLYGKRYIEFDF 191
 QY 186 XKIBEKYEKQEWPFVLYILIRKERYIKTSNVIEVEKEQIAKEKIIQKTLKEILYERP 245
 Db 192 SKIKEEYKQEWFIILFYLIRKERYIKTSNIIISVEKEQISKEKIIQKTLREILYERP 251
 QY 246 DDSIVFKIANKIYDKEFKLSCKEITHTLASILSESIIFILRIEFLAVIENDIDPKKILQ 305
 Db 252 DDSIVFKIANKIYDKEFKVSDKEITRHLASILSESIIFILRIEFLAVIENDIDPKKILE 311
 QY 306 ENKLYRSISPRYFFYDENTKKLEYKKIITIFNLLDKGSDAIKFPVPFNGGLFSBDKVY 365
 Db 312 ENKLYRSVSFRYFFYDENTKKLGYKKIITIFNLLDKGSDAIKFPFNGGLFAQDKVY 371
 QY 366 LNNEGLSISIEBIEILVMQLFFEEKNIKDEKFKVYSRLDPKSGELYETLLAYDIRIADT 425
 Db 372 LNNEGLSISIEBIEILVMQLFFEEKNIKDEKFKVYSRLDPKSGELYETLLAYDIRIADT 431
 QY 426 TVHRIEDGVVLTIRTEBELENKVNKIATYLGKNIYLTSSLDKSGGAYTTPDLDTPM 485
 Db 432 TVHRIEDGVVLTIRTEBELENKVNKIATYLGKNIYLTSSLDKSGGAYTTPDLDTPM 491
 QY 486 VISSIEEQLTKSPDLKIIDNSCGSGHFLISCLDYLTKVWYELDKFEDVKELDEEYR 545
 Db 492 VISSIEEQLTKSPDLKIIDNSCGSGHFLISCLDYLTKVWYELDKFEDVKELDEYVG 551
 QY 546 VIIIESEYDVQDSISKELVLRMLLKXCIYGVNDINPISVEITWLSLWINTFIPTPLSF 605
 Db 552 IILKESBEYDQDSISKELVLRMLLKXCIYGVNDINPISVEITWLSLWINTFIPTPLSF 611
 QY 606 IEHHIKAGNALLGYTKDEFFDVVKKESFSGSLFKKRIKEIITILEDIYOKIKGINDTTK 665
 Db 612 IEHHIKAGNALLGYTKDEFFDVVKKESFSGSLFKKRIKEIITILEDIYOKIKGINDTTK 671
 QY 666 EDIEKSKKIYKEYESKDIIDLRIIFSLIKLYSLSPDKSLNMFSDIASVISLIENILGN 725
 Db 672 EDIERSKKIYKEYESKDIIDLRIIFSLIKLYSLSPDKSLNMFSDIASVISLIENILGN 731
 QY 726 KTSSEDEKIEKIRKLSYKFFHYGIEFPDIOBGFQDIVIGNPPWEKTKFNTEFPSKHI 785
 Db 732 KISSDNEKMEKIRKLSHYKFFHYGIEFPDIOBGFQDIVIGNPPWEKTKFNTEFPSKHI 791
 QY 786 PNYRKLGIKEQNIIOBILSKONHPLSIEYNEEKNSIIANNIYKPKFCFTSGGPNLIF 845
 Db 792 PSYRKLSIEQNKIKQEIILGRNDHPLNTEYNEEKNSMGITNNLYKSDPFNFASGDPNLIF 851
 QY 846 RYFVTFNLKIKEKGNLTYLPSAIWNSSSRILRKHI FARYKLNLYIYQENKQRKPKOVH 905

Db 852 RYFAFNLKIKENGLNLYVPSALWSSSARILRKYIFTNKGNYIYQFQNRKPDVA 911
Qy 906 SSFKEAIFQLSNIKESTSFRAKFMIOSSDNILKEITRDLDKSDDAYKGIELNINQIKK 965
Db 912 TLFKFAIFQISNTKPTSFNFAKFMIOSSDNILKEITRDLDKSDDAYKGIELNINQIKK 971
Qy 966 LSPQIESIEFKDNEFTLINQMPKSPALGBGYIDFKKGLDPSIKNRKSLKKECNQVL 1025
Db 972 LSPQIESIEFKDNEFTLINQMPKSPALGBGYIDFKKGLDPSIKNRKSLKKECNQVL 1029
Qy 1026 IFLYSGANIHOVSFFEDKAKESKLLWIDKEDLVKDKONQYQTFERYFRAISNT 1085
Db 1030 IFLYSGANIHOVSFFEDKAKESKLLWIDKEDLVKDKONQYQTFERYFRAISNT 1089
Qy 1086 NERTMISTLSPGNCYCVANSIYINDKTPISLYKCLFIISFNSFVDFLLRRF-VDSNVL 1144
Db 1090 DIRTWISTLSPGNCYCVANSIYINDKTPISLYKCLFIISFNSFVDFLLRRF-VDSNVL 1149
Qy 1145 KSLYQCPMPQPEEKESKLLWIDKEDLVKDKONQYQTFERYFRAISNT 1204
Db 1150 KSLYQCPMPQPEEKESKLLWIDKEDLVKDKONQYQTFERYFRAISNT 1209
Qy 1205 ILKLDKEDEFFKEKENENNFIASLYSLAKEDPITLLGDPKALNKKKGEDYISSLIKGY 1264
Db 1210 ILKLDKEDEFFKEKENENNFIASLYSLAKEDPITLLGDPKALNKKKGEDYISSLIKGY 1268
Qy 1265 DNYLNNKI 1273
Db 1269 ENYLNK 1277

RESULT 3
E81339
probable restriction/modification enzyme Cj0690c [imported] - Campylobacter jejuni (str
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: E81339
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
R.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: E81339
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1250 <PAR>
A:Cross-references: UNIPROT:Q9PPL7; UNIPARC:UPI00000C1CEE; GB:AL139076; GB:AL111168; NID
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj0690c

Query Match 27.2%; Score 1781; DB 2; Length 1250;
Best Local Similarity 34.7%; Pred. No. 1.2e-60;
Matches 457; Conservative 244; Mismatches 479; Indels 138; Gaps 40;

Qy 5 DIVTNNPNISLYKOLSKDFIKENINLKDPF-----ILIKNKLFSIDDDNSTEAN 55
Db 6 EKKYLNINLFTSYSLIEDFPKINYNFANQKAPRDLITKTIYDNKPIKQNEHQFE-- 63

Qy 56 IESLLKYPPEELNYSVEQKAGIEGVSERVDILLFENDKOKASFNNKLKEAKQNEPTP 115
Db 64 -DEFISKVLEILGWCFVRQDEKIIQKLEKDPFLFNSNDKLKSKYENLDKTKSS- 118

Qy 116 IEDILLIARVPTSFDAKOLKESD-OLRYLNQYQKHYGIIISNGKWRLYDKSKVL 174
Db 119 -NDFTILLESKAYNEIDNK-KVKONPHFQILRYLGNLKKNGYGLTNGRFRFYDNS-IL 175

Qy 175 YGEKRYIBENFYKIEKEVEKQEWFLVIYILIRKERYLKTNSVLEVEKEQIAKEKIIQ 234
Db 176 NSGNVFIENLEKIIEDQNI---EAFAYFYSVAFNFTKEKEDHLEITLQNNKLSKIE 232

Qy 235 KTLKEILYERPDSDIVFK-IAKNYDKEPKLSGKEITQHILASILEESIIFILRIFFIAY 293

RESULT 4

Db 233 DDLKSIYGTNGNESLEFFTSGRIYNK-----TKADLKLYENSLEYIFRLLFIAY 283
Qy 294 IEDNDIFKKILQENKLYRSSISFRYFF--YDENTKCKLEYKKIITIFNLLDKGSDAIKPP 351
Db 284 PEDK--FEILLEKHFKYSKISLRTLENLQDESSSGGFELENI FNIYKNGKGNPOMP 341
Qy 352 VFNGLSPEDSKVYLNNEGLLSISEIBELVQMLFFPEKNIKDEKFKVYSLRDPKSPFEL 411
Db 342 VFNGLSPEDSKTALLTPKIFNDKDLKFLNQLNFKDNLSPFR--DYKTLSEVHLGTI 399
Qy 412 YETLEVDLRADTTVHRIITEDGVYLIRTEELE-----NKKVANKIATYL 456
Db 400 YEGLLSTPFEITAN-----EDIYVYKESKIECYFDNYDFKILEKSKKVEKYTYK 452
Qy 457 KGNITLRSRDLRKRKSGAYTTPDLDTFPMVISSEEQKTKSPLDIKIIDNSCGSGHFLI 516
Db 453 KGOIYLNKTSNSRKASAFYTPQSIANFLIOALQKDLNANNILKFKILDNACSGHFLV 512
Qy 517 SCLDYLTEKVMYELDFEDVKELDEYRVIISESEYDVQD-SISKELVLRMLLKXCI 575
Db 513 GVLNAITHIVLSDFDHFTNLKELYEEKEENILNHIKDF-VQDYEVDESILKRLILKRII 571
Qy 576 YGVDINPISVEITMLSWINTFIETPLSFTEHHIKAGNALLGYTKDEFFDIVKKPESG 635
Db 572 YGVDLNPFSIELTKLSLWIDSFIETPLSFTEHHIKGNALINSNLSDFKDLIK--QNS 628
Qy 636 PSLEFKRIEITILEDIYQKIKGINDTTFKEDISKSKKIYKEYBESKDIDNLRIIFSUK 695
Db 629 SNLPTNSITQEFELQEVFEKLDNLKDTNEBQIKQSKQIYQN-EITPKLNKLNLVLYIN 687
Qy 696 LYSLSFDKSNMPSDIASVISLIENTGNKTSSEDEKIEKIRKLSYFFHYGIEFP 755
Db 688 --TLHF--VNKEELQILKALS--QDDIQNL-SONEQAKAI--ISKYQEFNFYFYLEFP 738
Qy 756 DIQE-----GFDIVGNPPWEKTFEFTEFFSKHIPNYRKLGIKEQNLIKOELISKDNHP 810
Db 739 EIVENQVFKGPDIIIGNPPWDKTFSDSDFPQYKSDYRSLSIASKCKEIQNLAKOY-- 796
Qy 811 LSIENVEKNSIIAINNIYKDFDKFTSGGPNLFRYFVTNMLKLIKEGNLTIVLPSAI 870
Db 797 IKQYKQKAYINDLSEYKAYPLNKGSGDGNLFLPEVKNSLLKQDGLAYVLPAL 856
Qy 871 WNESSRILRKHI FARYKLYNIYOFENKCR-FKDVHSGFKFAIFQLSNKSSTSFKA 929
Db 857 MPEDGSLILRKEILENKTLEYFSPENNAIFIDVHRSYKPAALMIKN-TQANHHTKIKM 915
Qy 930 MIQSSD-NILKEITRDLDKSDDAYKGIELNINQIKLSPIQESIIIEKDNKEFTLINKM 988
Db 916 MFYKTDINSLK-----NKDEI--LTLNLKDIKKLSPTHLALMELKQALBILRKS 964
Qy 989 FSKFSALGEGYIDPKKGLDPSIKNRKSLKKECNKNLIFLYSGANIHOVSNSRF-----PE 1043
Db 965 YNAFQNSLFDYIDFRELDD--MTNDKDLFIEFREGGLPLYEGKMIHOFDANFSQATYFL 1022
Qy 1044 DK--DAKESKLLWIDKEDLEKVLIT-KDNOYQTE--RVFYRAIASNTNERTMTSLSPG 1097
Db 1023 EKAKFDEKLSKELYRACKATGKELNPKLIKVDREFFELGYRKIASDTNERTLIASLLPK 1082
Qy 1098 NCYCVNSIYINDEK-----TPISLYKK-LFIIISFNSFVDFLLRFRPVDNVLKS 1146
Db 1083 NCGGADSPYSNIPQYVYLKDDVICMDIIVPERILFVLALFNSLVVDFIIRNMVQINVS 1142
Qy 1147 CLYQCPMPQPEEKESKLLWIDKEDLVKDKONQYQTFERYFRAISNT 1203
Db 1143 YLERILPQPEDEEIQONNEIYKTLAKNALLQLVNDQN-----RHFDLKLQFNKIN 1194
Qy 1204 -KILKLDKEDFFKEKENENNFIASLYSLAKEDPITLLGDPKALNKKKGEDYISSL 1260
Db 1195 BEIPKTKAYDILRAK---NDLLVKELYGLSDDEFSYMSIFKVLNEKQ--SEYITLL 1247

D70251
 Hypothetical protein BBK02 - Lyme disease spirochete plasmid K/lp36
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
 R:Accession: D70251
 R:Praser, C.M.; Caejens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
 son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
 ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
 Nature 390, 580-586, 1997
 A:Authors: Smith, H.O.; Venter, J.C.
 A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
 A:Reference number: A70100; MUID:98065943; PMID:9403685
 A:Accession: D70251
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-422 <KLE>
 A:Cross-references: UNIPROT:050806; UNIPARC:UPI00000568A0; GB:AE000788; NID:g2690123; PI
 A:Experimental source: strain B31
 C:Genetics:
 A:Genome: plasmid

Query Match 25.2%; Score 1648.5; DB 2; Length 422;
 Best Local Similarity 75.7%; Pred. No. 3.7e-56;
 Matches 315; Conservative 49; Mismatches 49; Indels 3; Gaps 2;
 QY 853 LKLIKELNLTLPVSAIWNSSRIIRKHIFARYKLNIIYQFNKQKQFVHSSFKPAI 912
 DB 1 MKLVKPGNLTTLTSSLSSESGSKALRQHIIFSNKLNIIYQFNKQKQFVHSSFKPAI 60
 QY 913 FQLSNIKESTSTFKAFMIQSSDNILKEITRLDKSDKDAYKGIENINQIKLSPIOES 972
 DB 61 FQISNSKEPTFRVKFVIQSSDNIMKEITSDLKEGNEYAKGIELDIAIKRLSPIOES 120
 QY 973 IIEFKDNEFTLQNKWFSFSGALGEGYIDPKGLDPSIKNRKSLKKECNKNLILFLYSGA 1032
 DB 121 IIEFRDSAEFTLVNKFSPFDTLQIYIDFREG--NLTKYKALYKNEYNEKFIPLYSGA 178
 QY 1033 NIHQFSRPFEDKDAKESKLLWIDKEDLEKVLTKDNOYQTRVFYRAIASNTNERTWIS 1092
 DB 179 NIHQFSRPFEDKDAKESKLLWIDKEDLEKVLTKDNOYQTRVFYRAIASNTNERTWIS 238
 QY 1093 TILSPGNCYCNSIYINDEKTPISLYKKLFIISFNFSFVDFLLRFRVDSNVLSKLIYQCP 1152
 DB 239 TILSPGNCYCNSIYINVEIPISLYKKLFIISFNFSFVDFIIRFVNSVLKSLIYQCP 298
 QY 1153 MPQPEKEIILNPLYLNIAKNTSLIAKNDPENFKYLLYLFYFKDKEKKNKILKDKED 1212
 DB 299 MPQPEDEILNLSYLTAKNTSLIAKNDPENFKYLLYLFYFKDKEKKNKILKDKED 358
 QY 1213 EFPKKEKENNFIIASLYSLAKEDFTLLGDFKALKNKKKGEDYISLLIKGYDNYL 1268
 DB 359 EFPKKEKENNFIVASLYSLTKEDFTLLNDFKVCN--KKGEDYISLLIKGYENYL 413

RESULT 5
 T08313
 conserved hypothetical protein H1130 - Halobacterium sp. (strain NRC-1) plasmid pNRC100
 C:Species: Halobacterium sp.
 A:Variety: strain NRC-1
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
 C:Accession: T08313
 R:Ng, W.V.; Ciufo, S.A.; Smith, T.M.; Bumgarner, R.E.; Baskin, D.; Faust, J.; Hall, B.;
 Genome Res. 8, 1131-1141, 1998
 A:Title: Snapshot of a large dynamic replicon in a halophilic Archaeon: megaplasmid or m
 A:Reference number: Z16408; MUID:99063795; PMID:9847077
 A:Accession: T08313
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-706 <DAS>
 A:Cross-references: UNIPROT:052000; UNIPARC:UPI00000631E7; EMBL:AF016485; NID:g2822278;
 A:Experimental source: strain NRC-1
 C:Genetics:
 A:Gene: HALOSP:H1130

A:Genome: plasmid pNRC100
 Query Match 8.5%; Score 558.5; DB 2; Length 706;
 Best Local Similarity 24.1%; Pred. No. 1.9e-14;
 Matches 180; Conservative 150; Mismatches 219; Indels 199; Gaps 32;
 QY 62 YIFELNYSVQQKAGQIEGVESVDIL-LPENEKDKASFNKKLKEAKKNNEP-----I 114
 DB 18 YLDEHLPETEWEDEVSEELQAYDDIMALWEREKDTAPKRNESQLEBEKFIKPRKGLI 77
 QY 115 PIEDILIAE-----VKRPTFSP-----DAKDKLKESED----- 143
 DB 78 PFE-----VEESTSRTPRDYGFDTEDADAFEREEGGDFYENAVADAKRWGRPL 133
 QY 144 -----QLVRYLNQYKHGYGILNGKVRVLYDKSKVLYGE-----KRYI 181
 DB 134 DTRSGEHERDFENPSQIHVLOETPARWAVLTDGKKRWLY-----YGTSHRLDSYY 187
 QY 182 EFNFKIIEKEEYKEQEWFLVIYILIRKERYLKTNSVIEVEKEQIAKEKIIQKTLKEIL 241
 DB 188 EVDLPTILESGLDEDFKYFLF---FRHEAFLDGSG-DCFLLDDVYDSNVFAQELGEDL 243
 QY 242 YERDDSTIVKIAKNIYDKFKLSKEITQHLILASILEESIIFILRIFPIAYIENDIFPK 301
 DB 244 QDN-----IYEAIKILSEGLQYPPENDLDEGLDILIHDSLSIYLYRFLIYVAAEAG--R 296
 QY 302 KILQ-ENKLYRSSISF-----RVFFYDENTKKLEYKKIITIFNLLDKG 344
 DB 297 ELLDTSNIYFQSYSLNSLKKQIAEELDSGPKYRDMQDNLOQSLD-----ELFTLINK 351
 QY 345 S-----DAIKFPVFNGLFSED-----KVLYLANE--GLISISIEIEILVKMLPFE 388
 DB 352 SKSRGIPEDLIYIPAYNGGLFRTDDEDDSAEAKELANHDVGDALVAKVELLTR----S 407
 QY 389 EKNIDKEK-FVKYGRLDPKSGELYETLLELYDLRIADTVTHRIIDGYVILRTBELENK 447
 DB 408 KNDGGGKIFVDYSSLDVRHLGSIYEGLLYQLANVADEPL--ALDDGSYV--SADEGDDV 463
 QY 448 KVKIATVYKNIYLTSLRDLKSKGAYTTPDDLTDNV-----ISSIEEQLATKSP- 499
 DB 464 VVQ-----EGEVLTDTDSGERKATGSIYTPYVVEIVENTLEPLVDIDRADLAGRSAR 517
 QY 500 -----LDIKIIDNSGSGHFLISCLDYLTEKVMYELDKFEDVKKELDEBYR 545
 DB 518 GEDRFAFEAEAFRIFDLKILDPAMSGHFLTSAIDYLAREI-----IDAQER 564
 QY 546 VIIEB-----SEEDYVDQDSISKELVKNMLKXCIYGVNDINPISVEITWLSWINTFI 600
 DB 565 QAAQQGIETVQDHDINWA-----RRQVQRCCIYGVDLNPLAVELAKVSLMLRTAAE 617
 QY 601 TPLSFIEHHIKAGNALLGYTKDEPFDIYKKK-----PESGFSLPKRIKEIT 648
 DB 618 QPLAFDLHLTKTGNLSVGSNVEE-IDALERNGAABEGPNSTLIDFGMT-FEGTWENLMA 675
 QY 649 ILEDIYQKIKGINDTTKEDIEKSKYIK 676
 DB 676 IHADLI-----ALENEELEDVKAMEKKYQ 699

RESULT 6
 B64516
 hypothetical protein MJEC502 - Methanococcus jannaschii plasmid pURB801
 C:Species: Methanococcus jannaschii
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
 ; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.,
 rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese,
 A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
 A:Reference number: A64300; MUID:96337999; PMID:8688087
 A:Accession: B64516

A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1181 <BUL>
 A:Cross-references: UNIPROT:Q60301; UNIPARC:UPI000013C258; GB:L77119; NID:g1500688; TIGR
 C:Genetics:
 A:Map position: ECSREV4814-1269
 A:Genome: plasmid
 A:Start codon: GTG
 A>Note: this stable 16-kilobase pair plasmid is also designated ECS (small extrachromosome)

Query Match 8.1%; Score 526.5; DB 2; Length 1181;
 Best Local Similarity 22.3%; Pred. No. 5.8e-13;
 Matches 322; Conservative 212; Mismatches 399; Indels 509; Gaps 76;

QY 25 IKENINKLQPFILIKVLSIDD-NSTEANIESLLK--YIFELNYSVEQKAGQI- 79
 DB 15 IPKEVYNKQFQIKLVGNILSIKQIPLEGKFKELLKGRYFKSDDF-----GGQLP 68
 QY 80 -----EGVESRVDILLFENDKDKASFNKL-KEAKKNNEP-----IEDILIAEV 125
 DB 69 BPFTQKVIPIEPLGYE-----FTSEIKSKSPGLGRKIPDYRVSVFNKEILIEAE- 120
 QY 126 KRTPS-PAKD-----KLKESDQLYRYLNOYQKHYSILSGKWRLYDKSVLYGKRY 180
 DB 121 -----PLGSLNKKDSGHQVKE-----WLIKSYGVDGTGATNGLEWVLLHYDDTIKEITL 173
 QY 181 IEFNFKIEE-----KEEYKEQEW---FVLPIYLIRKERYLKTNSVIEKEQIA-KEK 230
 DB 174 KELNLSIFVVLNKKDLLENELKQVSEFYCFSEK-YIE-EYIEVATKNIKHKE 230
 QY 231 EIIQKTLKELLYERDDISIV-FKIAKNIYKFKLSKEI-TQHILASILEE----- 280
 DB 231 EITNEFYKEFV-----KLVGFDVQVKKQSSSEKQKGTCKCLYNCIEAPNTSEL 284
 QY 281 -----SILFILRPIFIATED-----NDIFKKILQENLYRSSISFRYP-----FYD- 322
 DB 285 DKKFAVLLMRLFIKLEKXGIVPRLRLRTYEDYK--KSNVLINYYDAYLPLFYEV 342
 QY 323 -----ENTKKLEYKKIITIFNLLDKGSDAIKFPVNGGLFSDKVKYLNNEGLLS 373
 DB 343 LNTPEDERKENIRTPYKDI-----PYLNGGLFRSNNVP-----NELSFT 383
 QY 374 ISEIEILVKMLFPEEK---NIKDEKFKVYSLRDPKSGELVETLLEVDLRIADTTVRI 430
 DB 384 IKD-NEIIEVINFLERYKFTLSTSEGSEVELNPDLGYVE----- 425
 QY 431 IEDGYVLRTEELENNKVKNIATYKNIYILTSRLDKSKSGAYTTPDDLTDNFVSSI 490
 DB 426 -----KUNILAE--KG-----QKGLGAYTTPDEITSYAKNTI 457
 QY 491 E-----EQLK-----TKSPLD-----IKIIDNSCG 510
 DB 458 EPIVVERFKEIKKNWKINDINFSTLDELNEDSKIAENKHILRAFLDELDKIRILDPAGV 517
 QY 511 SCHFLISCLDYITEKWTYELQKFDVKKELDEYRVIIIESEYDVQDSISKELVKRLM 570
 DB 518 SCHFLISALKELLQ-----IKKRI-----YLLREEMDIY-----KE---KLGI 553
 QY 571 LKXCIYGVNDINPISVITWLSMI-----NTPIFGTP-----LSFIEHIKAGNALLGYTK 621
 DB 554 ILNNLYGVNDIDIAIEIAKLRLWALINLDEALKRGVELLPNIEYVNRGNSLWIMID 613
 QY 622 DEFFDVKKFSGSLFKRIKEITITILEDIYQIKINDTTEDEIKSKKIYKEYEES 681
 DB 614 ENL-----KQLSISYLCNVR---IMCVLEGLI-----INAHNSERKKLKA-KELLEK 659
 QY 682 XD-----IDNLRIFSLI-KLYLSLSPDKSLNM-----EFSDIASVISLIENTL----- 723
 DB 660 RDGYVLDNVVEAYHLLYEVYRTHGKANLLKELLDEIRD-----SIVESTVPAYFABIY 714
 QY 724 --GN-----KTSSEDEKIEKIRKLSYKFFHYGIEFPDI--OEGFDVIGNPPWKT 773
 DB 715 QNGNNKONGKSKKNRPRVEFEKL-----KPFHWKIDFGWIIKEEGFDVIIGNPPYGNL 770

QY 774 KFNTEPFSSKHIPNRYKLGIKIQEONIIQOELSKDNHPLSIEYNEEKNSIIAINNIYKPDF 833
 DB 771 -LSPT-----KEIMKRDTP-----EFD- 788
 QY 834 KCFTSGGDPNLFYVFTFNKLIKEKGNLTVLVPSAIWNSSSRILRKHFARY---KLN 890
 DB 789 -----IFVTIVHSSKLLKNEGVLGFIIPSSFGTGVRYNLRKELFTKCKLKI 838
 QY 891 Y-----IYOFENKGRKFDVHSSFKFAIFOLSNIKE 920
 DB 839 YLPEDVSGAYVDNIIILHKPKPSBDLVLIYAPPKTK----- 878
 QY 921 STSSFKAKFMIOSSDNILKEITRDLKSKDAY-KGIELNINOIKULSPIQESIIIEPKDN 979
 DB 879 -----KISFEKNDLFIYSKILNDPKCRIFPKSPFIYIILDKIKONCRESUTYLEDL 931
 QY 980 EEFLLINKMPSKF---SALGEGYIDFKG-----LDPSIKNRKSLKCKCNKNILPLY 1029
 DB 932 TEST-IGILASKYFSEKENEYLYPLEGVNRYEYFKLKNVYDFSKHKNNEKLINLF 990
 QY 1030 SGANIHFNSRFFEDKAKESSKLLWIDKEDLEKVLTKDNQYQTERVYFRAIASNTNERT 1089
 DB 991 -----MSPEKIFIRRIV-NRQDRI 1008
 QY 1090 MISTLSPGNCYCVNS-----IYNDEKTPISYKKLFIISIFNSVFDLLRRFVDSN--V 1143
 DB 1009 MASY---GNIBGVYKDYLVFLKPDTPINFY---YLLGILNSELISYI---YIGKSAIA 1059
 QY 1144 LKSCLYQCPMPQPEKEILSNPLYNLAKNTSLIAKNDPENFKVLLVLYFKPDKEKNV 1203
 DB 1060 LKDDFQTTLEELRELPIV-----IPKNKIINA-----LTQLSKURFELD-- 1101
 QY 1204 KILKXDEDEFFKEKENENNFIASLYSLAKEDFITLLGDFKALKKKKGGYDISSLIKG 1263
 DB 1102 ---KLNENDRIFLE-----NIIDSLVYGIYFQDLIP-----KEELNEICNE--INGIILK 1146
 QY 1264 YD 1265
 DB 1147 YD 1148

RESULT 7

E70251

hypothetical protein BBK03 - Lyme disease spirochete plasmid K/lp36

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004

C:Accession: E70251

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
 son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
 Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
 Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A>Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943; PMID:9403685

A:Accession: E70251

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-134 <KLE>

A:Cross-references: UNIPROT:O50807; UNIPARC:UPI00000568A1; GB:AE000788; NID:g2690123; P1

A:Experimental source: strain B31

C:Genetics:

A:Genome: plasmid

Query Match 7.2%; Score 473; DB 2; Length 134;

Best Local Similarity 72.0%; Pred. No. 5.1e-12;

Matches 95; Conservative 18; Mismatches 19; Indels 0; Gaps 0;

QY 606 IEHHKAGNALLGYTKDEFFDVKKKYGSGSLFKRIKEITITILEDIYQIKINDTTEK 665

DB 2 IKHHKVGNALLGYTKDEFLNLFONEFRGNCLSVVVKINEITILEDIHOIKGINITYIK 61

QY 666 EDIEKSKKIYKEYEESKDIDNLRIFSLIKLYSLSFDSKSLNMFSDIASVISLIENTLGN 725

Db 62 EDLEKSKYKEYQENINNLRIIFSLIKLYSLFSDKSLNIKFSDITVINLIGNILDS 121
 Qy 726 KTSSEDKKIEK 737
 Db 122 KISSEDKKIEK 133

RESULT 8
 D71810
 probable type II DNA modification enzyme (methyltransferase) - Helicobacter pylori (stra
 C:Species: Helicobacter pylori
 A:Variety: strain J99
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
 C:Accession: D71810
 R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
 Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jlang, Q.; Taylor, D.E.; Vovis, G.F.;
 Nature 397, 176-180, 1999
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
 A:Reference number: A71800; MUID:99120557; PMID:9923682
 A:Accession: D71810
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1252 <ARN>
 A:Cross-References: UNIPROT:Q9ZJA4; UNIPARC:UPI00000D373F; GB:AE001563; GB:AE001439; NID
 A:Experimental source: strain J99
 C:Genetics:
 A:Gene: jhp1409

Query Match 7.0%; Score 455.5; DB 2; Length 1252;
 Best Local Similarity 20.7%; Pred. No. 3.1e-10;
 Matches 302; Conservative 202; Mismatches 499; Indels 453; Gaps 61;

Qy 22 KDFIKEN-----INLKDFILIKNKLFSDDNSTEANIESILK--YIPELNYS 70
 Db 9 KVFIIKYNPOEPKGTIENFEKEINSLENVPRQDDEEFQKNEINKPLKNAYGYDCNTY- 67
 Qy 71 VEQKAGQLEGVESRDILLFENDKDKAFNNKLEAKKNNEPIP-----IEDI- 119
 Db 68 ----EKVDSVIYDKEVRVLI-----EVKALNNKTEFPKQRENPISKACQWVLYFLEEIE 119
 Qy 120 -----LIIAEKVRPTFSDAKDKLKESED-QLYRYLNOYQKHGYGLNSGKWRLYD 169
 Db 120 KEKNNSLKHTIICNA-HBFFLFDCKDLSLKEDKRIKKFYKNYAKKEGT-----D 168
 Qy 170 KSKVLY-----GEKRYTEFPNFKIEEYKQWVLYFYLRKERYLKTS 216
 Db 169 SSKPKFYKDLQEQDFQGLRYTHFNJ-----SDDFKE--LPLIYQVLSQEVLLKQK 220
 Qy 217 NVIEVEKEQIAKEKEIOKTLKEILVERPDD--SIVFKIA--KNIVDKFELSGKEITQH 272
 Db 221 RYLDAN-----TLNKDFYEBLLIYLGEBQNDKGKILIKPSRTNSLSGALKKYNKLQDE 276
 Qy 273 ILASILEESIIFILRIFFIATIEDNDIFKKILOENKLYRSSISFRYF-----FYDENTK 326
 Db 277 EWMALL---IANNRIFLRLESLLISPEHPEKERSFLTTGNFKDFNALNTLFEVLAK 333
 Qy 327 KLEVKYKIITIYNLIDKGSDAIKFPVFNGLPSE-----DKVKYLANEGLLSISEIEE 379
 Db 334 KNSRLPEIKENKILG-----KIPYNSSLFDKPTPLKLGHEIKLLDNKKL-----E 380
 Qy 380 ILVKMLFPEEKNIKDEKFKVYSRLDPKSPGELYETLLVEYDLRIAD-----TVHRIEDGV 435
 Db 381 IYNSVLKKHENYQDKK-----ASLPLELYFPKFLHLYKFTTPDKDKONT 426
 Qy 436 YLIRTEEELENKVNKIATYLNKGNITLRSRDLRKSQGYATPDBLTDPMWISSIE----- 491
 Db 427 --DTSESRLINFSV-----LGLVPEKLVNGYKESGYFTSPFITSVMCKESITIVL 474
 Qy 492 -----FQLKT-----KSPDLKILDNCSGSGHPLISCLDYL 522
 Db 475 DKFNQKYNIECDKJELKNLYLNKSNYKEDKKRYQLQLLTLCRDPAGVSGHFLVSNALNEM 534

Qy 523 TEKVVYELDKFEDVKKELDEBYRVIIESE-----EYDVQDSISKE 563
 Db 535 V-LIAYELGLIASLHR-----YSLTENDEIIIHVTPGEIFNYKKPLRENDPHHHIQKE 588
 Qy 564 LV-LKRMLLKKCIYGVDPINPISVEITMLSLWI-----NTFIFGTPLSFIEH 608
 Db 589 LFEKLSIIENCLFGVDINPNSCBITKLRLWIELLKYSYIFIEGKNTNALET-LPNIDI 647
 Qy 609 HIKAGNALLGYTKDEFDDIVKKPFESGFSLPFKRIKEIITIL-----EDIY 654
 Db 648 NIKANSLI--SRFNLDLKK-----IPNIKKIQEYKDLVAQYKDPNPYPLNKADLI 700
 Qy 655 QKIKGINDT-----TKEDIEKS-KKIYKYEESKDIDNLRIRIIFSLIKLYSLSPDKS 704
 Db 701 NKIQDLKNTFSLTLKDPKTKAELEKTEKHKKYN-----FPAL-----DDKS 743
 Qy 705 LNMESDIASVISLIEINLGN-KTSSDKE-----KIEKIRKL-----SSYKPFH 749
 Db 744 L-----LDGLNYPFIPSLFGLTKLSPKEEBEAFASYGRIRALRKKLDDALSGGEYQNAFE 797
 Qy 750 YGIEFPDIOE-----GFDIVIGNPPWEKTKFNETEFPFSKHI PNVRKLGIKEQNIKQEI 803
 Db 798 WRFEFPEVLDDGDFLGFDCIIGNPPY-----IROEQ 829
 Qy 804 LSKDNHP--LSIEYNEEKNSIIAIINNIYKDFPKCTSGDPPNLFYFVFNKLKIKEGNL 862
 Db 830 I-KDKLPLEKQYQDFYNS-----TADIYTFPALSFPHLLEKGF 869
 Qy 863 TYLVPSAINNESSRIILRKHIFARYKLNYYIYQFNKRFKDVHSSFRPAIFOLSNIKEST 922
 Db 870 AFITSNYARAKYGAKLREMLKKTTLVSYMELNALK-----VFESAADVTSI 917
 Qy 923 SSRAKAFWIOSDNILKEITEDLAKS-KDDAYKGIELNINOIKKLSPIQESIIEPKDNEE 981
 Db 918 IHFIKQPLKESDFKYEPTDDDKDLKSTPYLSMKQNALSTESFIFANATLLDURD--- 974
 Qy 982 FTLINKMFSKFSALGEGYIDFKKGLDPSIK-----NRKSLLEK-----NN 1022
 Db 975 -----KIESVGTPLKWDIOINVCIGIKTGANEAFIIPTEKREILNACKTQBERET 1025
 Qy 1023 KNLI-----FLYSGANIHOFSRPFEDKAKESSKLLMIDKE-----D 1060
 Db 1026 ERLIKPILRGKDIKXYSYEWAGELINTHNGYTSNLK--SKIPPIDIEKYPATKAHLDAH 1083
 Qy 1061 LEKVLTK---DNQYQIYERVEYRAIASNTNERTWISTLSPGNCYCVNSIYINDEKTPISL 1116
 Db 1084 YDTIATRCQDGTTPYHLRNCAI--LEDPEKEKIVYPETSSQGAIFYENSGIFLEKTVFMI 1141
 Qy 1117 ---YKKLFIISFNSFVDFLLRRFVDSNVL-----KSLYQCMPQPEEKEILSN 1164
 Db 1142 VSDAYNLKLLTALLNSKLITFYFNFCGCGCILKSGYQYNKHALEKIPIPQITEKNQOLA 1201
 Qy 1165 PLYNLAKNTSLLIAKNDPENFKYLLYLEYFKDKEKVNKILKDKEDBEFFKEKENNPF 1224
 Db 1202 DKITDCAEQIILAKAKOPKANTOKL-----EKEIDA 1232
 Qy 1225 IIASIYSLAKEDFITL 1240
 Db 1233 LVYQLYHTDDEIKTI 1248

RESULT 9
 B81419
 probable type IIS restriction /modification enzyme, N-terminal half Cj0031 [imported]
 C:Species: Campylobacter jejuni
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
 C:Accession: B81419
 R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
 N.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell
 Nature 403, 665-668, 2000
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
 A:Reference number: A81250; MUID:20150912; PMID:10688204
 A:Accession: B81419

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-867 <PAR>

A:CROSS-references: UNIPROT:G9PJ080; UNIPARC:UPI000000C2009; GB:AL139074; GB:AL111168; NID

A:Experimental source: serotype O2, strain NCTC 11168

A:Genetics:

C:Gene: CJ0031

Query Match 6.8%; Score 446; DB 2; Length 867;

Best Local Similarity 22.4%; Pred. No. 4.7e-10;

Matches 228; Conservative 141; Mismatches 291; Indels 358; Gaps 45;

QY 34 KDFPILIKNKL-SSIDNTEANIESLLKIFELNYSVBOQKAGQIEGVSRVDILL 90

DB 9 KDFP-PPYRKQIQMONEFIDISKALMQL-ERLESSQSENEEDYLVANALSPFLTWL 63

QY 91 FENDKDKASFNKJKEAKNNPIPI-EDILIAEVKRP-TSPDAKDKLRES 141

DB 64 -NFKTHIKTKQKSEIDLISKDFSKDLVLIEAKKPNSEKFEITHKVNK 115

QY 142 -EOLYRLNQYKHGILSNGKWRLYDKSVLYGKRVIEFNFXKIEKEE--YKE 196

DB 116 ALHETILYFRN-----REYSPSLFIIITDFYKFYIFKISFEELFYKN 160

QY 197 QEMFVLFYLIR-----KERYLKTNSVIEVEKEI-----AKEKE- 231

DB 161 PSFKKLFEFCNPNSLFKNTTEFYKVAKLNSKENLKGFLIDLTFLLKQKGNPKNL 220

QY 232 -IIQKTL-----KEI-----LYERPDDSIKFIKNIYDKI- 261

DB 221 ASIYKTFHRDLSSEFNPNANSLNNAFYKELLYILGLYESKONKLI-IAKSESEKEQ 279

QY 262 -FKLSGKEITQHILASILESIIPILRPIPIAVIEDNDI-----FKKILQE 306

DB 280 GTFYTAINSKLEENFETIKLLILNARILFKLIESNLVFNDDKNLFLNFKKIPDF 339

QY 307 NKLYRSSISFRFFDENTKKLYKIIITINLLDKGSDAIKFPVFNGLSEDKVKYL 366

DB 340 DKL--SELFVLAKESTRKSEFTYL-----PYNLSLFKOSI--- 378

QY 367 NNEGLSISEIEILVMGLFFBEKNIKDEKFKVYKSLRDPKSPGELVETLLVDLRIADTT 426

DB 379 -ENTLISLNDL-KLFYKNTVLKDK-CKAKGQVGLLEYLFEFLDSFDGSD- 432

QY 427 VRIIEDGVYLIRTEELENKVNKIATYKGNIVLTSRLDRKSGAYYITPDDLTDFW 486

DB 433 -----EQSEILSQEL--ISSSVLGNVF--EKLNGYKESGYTPSPITSYMC 475

QY 487 ISSI-----EQLTKSPL--DIKIDNSCGSHFLIS 517

DB 476 KESITKVLDFKNAQFDLDAKDISLRKSLRDKKAKQKELLNSIKICDPVAGSGHFLVS 535

QY 518 CLDYLTEKVMYLDKFEDVKKLDERYVILIESE-----EY-----DVQDS 559

DB 536 ALNVM-L-SYDENLF-----DEEYLEVQNDLEILITGRKGEFIEYKRPSTPKDTHL 587

QY 560 ISKELV-LKRMLLKXCIYGVDPINPISVETMLSLWINTPIFGTPLSF----- 605

DB 588 IQQELFHTKKDIENNLFGVDINPNSCEITKLRLMIELLKHSFYQSFDGNGVHDLKTLPN 647

QY 606 IEHHKAGNALLGYTKDEPFDIVKKFPGSGFLFKRIKEIITILEDIYQKIKGINDTTK 665

DB 648 IDINIKCGNSLSVY-----FETGKSL-----NHPNI-----K 675

QY 666 EDIEKSKYIYKEE-----SKIDNRIATIF-----SLIKLYSL 699

DB 676 ERMGYKRVLDYKEGYTKSHINOIKNLKISFKNFCFADKFKKEMKSFNDKCEKYSK 735

QY 700 SPDKSLNMFSDIASVISLIENTLGNKT--SSEDEKIEKIRKLSYYKF-----F 748

DB 736 KYGNFLAVDENLKPVS-----ANLTLFDFDEKEAIKEPANLKEVDNIENLSNHPF 789

QY 749 HYGIEFPDIQ-----GFDIVGNPPWEKTKFNETEPFSKHPNYRKLGIKEQNIKQE 802

DB 790 EWRFEFFELDDGNGFKGFLIIGNPPY-IRQE 821

QY 803 ILSKDNHPLSIYNEEKNSIIAINIYKDFKCFSTGSDPNLFRYVTFNKLKIEKG 860

DB 822 -----ELKELKPLAKNYKVYK-----GTSIDIYTYFELGFLVNLKDRG 859

RESULT 10

E64709

Type IIS restriction enzyme R and M protein - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004

C:Accession: E64709

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;

Peterson, S.; Lottus, B.; Richardson, D.; Dodson, H.G.; Glodek, A.; McKenna

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Matthey, L.

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.

A>Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: E64709

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1279 <COM>

A:CROSS-references: UNIPROT:O26046; UNIPARC:UPI000000D30D2; GB:AE000649; GB:AE000511; NID

Query Match

Best Local Similarity 6.4%; Score 419.5; DB 2; Length 1279;

Matches 308; Conservative 209; Mismatches 454; Indels 487; Gaps 70;

QY 68 NYSVEQKAGQIEGVSRVDILLPDKD-----KASFNKLKEAKK----- 109

DB 12 NYPNQ-----QKAPETAFDALETNQENENHQNDAFNLLKGFKYKVPFKKIDSTI 67

QY 110 --NNEPIPIEDILIAEVKRPSTFSAKD-KLKESEDQLYYL-----NQYQKHGYL 159

DB 68 LNENNE-----VEVIEFKAKNPNEFIKGGDLNVKAFHESLSLYLTERKEGNNLKL-IL 123

QY 160 SNGKVMRLYDKSV-LYGEKRYIEFNFXKIEKEEYKEQEWFLFIY-----LIRKERY 212

DB 124 ATIKELYIIDANEPEVFNKDEIENAFKNCHDR---KGNDRTRKAFYDACQRLNEFDRS 180

QY 213 LKTSNVIEVEKEQIAKEKIIQKTLKEILYERPDSDIVFKIANKIYKFEKLSGKE----- 268

DB 181 LK-VHYIPLKKNLAL--IYQALSPNLLKIPKYSNDANTLNKOFYELLTYLGLSEQND 236

QY 269 -----ITQHTLASIL-----EESIIFIL-----RIFFATYIEDNDIFKKILQE 306

DB 237 KKKLIKPSRTQNSLDALKEKYNLDDEEVWALLIANNRILFLRLLES----- 286

QY 307 NKLYRSSISFRFFFYDEN-----TKKLEYKKIITIF--NLLDKGSDAI-----KF 350

DB 287 -----LLISFKHF--ENPFLTENFENFDNLTLFFVLAKNSERLPEIKEDKILEKI 338

QY 351 PVFNGLFSE-----DKVKYLNNEGLSISEIEIILVMKLFEEKNIKDKK--FVKY 400

DB 339 PYLNSSLFDKPTLEKLGHEIKLLDNKCL-----EYKNSVLKHKDQYQEKPLPLLY 391

QY 401 SRLDPKSGELVETLLEVDLRIADTTVHRIIEDGVYLIRTEELENKVNKIATYKGN 460

DB 392 -----LKFRLRYKFP---TTTPKDKONT--DTSESRLINPSV-----L 425

QY 461 YLTSRLDRKSGAYYITPDDLTDFMVISIE-----EOLKT----- 496

DB 426 GLVFEKNGYKESGYTPSPITSYMKESITPVLDKFNALYQWDCENLKLARGEIDRNP 485

QY 497 -----KSPDLIKIIDNSCGSHFLISCLDYLTEKVMYELDKE-----DVKKGLDE 542

DB 486 SNEKAYELNTLTLRIKCPAVGSGHFLVSNLNMV-RVAVELGLIASLYRYDLKLENDE 544

QY 543 -----EYRVIIIESEEDVDQDSISKELV-LKRMLLKXCIYGVDPINPISVETMLSL 592

Db 545 IIIHHTPTGTEIFYNIYKPDSENDPHHHIQKELFNLKKSIIENCLFGVDINPNSCEITKLRL 604
Qy 593 WT-----NTPIFGTPLSFIEHHKAGNAL-----LGY 619
Db 605 WIELLKYSYIIEKKGKNTNALST-LPNIDINIKCANSLRSFALKDKKALLKSEKNLLEY 663
Qy 620 TKDEFFDIVKKFESGSLFK-KRIKEIIT-----ILEDIYOKI----- 657
Db 664 STAEYKELVK-----LYKPKILETLTHPIKDSNAVRKYAKERLYQELQNPKNOKP 715
Qy 658 KGINDTKEDIDSKKIYK--BYERSKOTDNRIIFSLIKLYSLSPDKSLN---MEFSDI 712
Db 716 KALNDR-----IEKIKAFKLTLPNPPKELFKFKFLKEHLELYCKSLBEANTNGLEAL 771
Qy 713 ASVISILNIGNKTSSDEKIEKIKLSYYKF-----PHYGLEF 754
Db 772 A-LEKOMANLFFDRYPKLDKSDKVVGLEHFNRYVLTYSYKLODENERYANALBWRFEF 830
Qy 755 POIQE-----GFDIVIGNPPWEKTKFNETEFPSKHIPNRYKLGIGKEQNIIEKQILSKDN 808
Db 831 PEVLDDGDFSGDCIIGNPPY-----IRQEH-KDL 861
Qy 809 HP-LSIEYNEEKNSIIANNIYKPDFKCTSGDGNLFRYFTVFNLIKIKERGNITYLVP 867
Db 862 KPILERKQYQDFYNS-----TADIYTFPALAFPHLLKEKGFSAFITS 902
Qy 868 SAINNESSRIILKHFAR-----YKLVYIYQENKKKPFKDVHSSFKFAIFQLSNIKES 921
Db 903 NKYTRAKYGAUKREWLLKKTITIVSYMELNALKVFSAA-----VDTSI-----IHFIKQT 952
Qy 922 TSSFRKFMIOQSDNLIKETRDLDKSDDAYKGIELNINOIKLSPIOESIIEFKDNBE 981
Db 953 PSK-----ESEFKYIETPNDKDLKSTPHLLMKQNVLTSTFIPANATILDLRD--- 1002
Qy 982 FTLINMFKSFGALSGYIDFKKGLDPSIKNRKSLLEKCNKNKILFLYSGANIHQFNRSF 1041
Db 1003 -----KIESVGTPDKWDIQUINGIK-----TGA-----NEAF 1030
Qy 1042 PEDDAKESKLLMIDKEDLEKVLTKDNQYQTERVPY-----RAIASNTNERTMTSLSP 1096
Db 1031 IIPTEKRE-----ETLNACKTOEERERETLIKPILRGKDTIKRYSYEWAWHLWVINT 1081
Qy 1097 GNCYCVNSIYINDEKPIISLYKKLFIISFNSVFDFLLRRFVDSNV---LKSCLYQCPM 1153
Db 1082 HNGYTS---LASKPIPIDIEKYPAKALDAH-YDTIATRCQDGTPHLRNCAYL--- 1134
Qy 1154 POPEEKE-----ILSNPLYNLAK-----NTSLLI----- 1178
Db 1135 -EDFEKEKIWASVGFVEYCMIPGLLILDTNVPFEVSKFGNTKNVLLGLNSKLLTFMLK 1193
Qy 1179 AKNDP-----ENFKY-LLYLEYFKF---DKEKNKILKLDKEDDEFFKEK----- 1218
Db 1194 AKNTPLGDMGVRNYKYNIMELPMWITAKNKKIADKIATLVDKILQAEKDKPANKTKL 1253
Qy 1219 ENENNPIIASLYSLAKED 1236
Db 1254 EXEIDALVYQLYHLTDEE 1271

RESULT 11
T28676
rhoptry protein - Plasmodium yoelii (fragment)
C:Species: Plasmodium yoelii
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T28676; A45521
R:Sinha, K.A.; Keen, J.K.; Ogun, S.A.; Holder, A.A.
Mol. Biochem. Parasitol. 76, 329-332, 1996
A:Title: Comparison of two members of a multigene family coding for high-molecular mass
A:Reference number: Z20507; MUID:97077455; PMID:8920022
A:Accession: T28676
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-2401 <SIN>

A:Cross-references: UNIPROT:Q26216; UNIPARC:UPI000017B647; EMBL:U36927; NID:g1041784; P1
R:Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.
Mol. Biochem. Parasitol. 42, 241-246, 1990
A:Title: Identification of the gene for a Plasmodium yoelii rhoptry protein. Multiple co
A:Reference number: A45521; MUID:91101660; PMID:2270106
A:Accession: A45521
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 2260-2401 <KEE>
A:Cross-references: UNIPARC:UPI000017B648; GB:M34281

Query Match 5.8%; Score 378.5; DB 2; Length 2401;
Best Local Similarity 20.0%; Pred. NO. 5.6e-07;
Matches 330; Conservative 247; Mismatches 527; Indels 547; Gaps 74;

Qy 14 ISLYKQLSKDPFKKENINKLXDF-----PILKNKLFSDDDNSTEANISSL-----LKLYFE 65
Db 253 VDLKKEIKEDKNVYIDELAKEPPYQITKYIEKKNEY-----NTIKSDFDKIYVGDIEQLYN 308
Qy 66 ELNYSVEQKAGQIEGVESRVDILLFENDKDKASPNKKLKEAKGNPEPIEDILIAEV 125
Db 309 EM-FSVVQES--SNIEHIENKTEILTKIDNVNINQMETE-----TV 350
Qy 126 KRPTFSFPAKDLKXESDQLYRYLNQYQKHYGILNSGVWRLYD---KSKVLYGE----- 177
Db 351 KSHLKNVETNNKLGSETILDIKIYI-----YGEITNELNKTLEDPKNEKGLSNKIDEYA 404
Qy 178 KRYTEFNFKK---TEEKEEYKEQWFVLYILIRKERYLKTNSVIEVEKEQIAKEILIQ 234
Db 405 KENVLYNKYSNILEIKKHYNQ-----INIDNIKEKAKQNYDQFKHEM 449
Qy 235 KTL--KEILYRPPDSDIVFKIAK-----NIYKEFKLSGK-----EITOHILA 275
Db 450 KIIPNEMKYOKP--SIEIKMKDEFKSKVKNYNDFKVYKEKVESEHNKFTELTNKTKT 507
Qy 276 SILEESIIFILRPIFIAYIEDNIPFKILQENKLYRSSISPRFYFFYDENTKKL--EYK 333
Db 508 EVSDEEI-----KK--YENKFNDKSLI-----NETKKSIEEYQN 541
Qy 334 IITFNLLDKGSDAIKFPVFNGLFS-----EDKVKYLNNEGLISIEIEILVKM 384
Db 542 I-----NTLKRVDDYIKVCLNTNELITNCHNKQTTLKDKL-----NQNKITKETNSI-DKI 592
Qy 385 LFFFEKNI-----KDEKFVKYSRLDPKSPGELYETLLE--YDLRIADTTVHRIIEQGVYLI 438
Db 593 YTDKFNILTDKKTELETFTKFTGLSLNNHESNNKELLTYFYDLK---ANLKKKENWLYKQ 649
Qy 439 RTEES-----LENKKNKIATYLNKGNIIYLTSLRSLDRKKSGAYYTPDDLTDFMV----- 486
Db 650 FNEKEKAVEDIKKQNVNDINKIVSNIEITITYSIYNEDTENEIGKSIELLNTKVKLEKVK 709
Qy 487 -----ISSIEBQLK-----TKSPLDIKIDNSCGSGHPLISCLDYLTE----- 524
Db 710 ANVTNLNBIKEKLKDYDFQDFKEKNIKYDPENKIKNDIDTLNQKIDKSIETLTIKQNS 769
Qy 525 -----KVVVELDK-----FEDVKKELEDEYRVITEESSEYVDQDSISKELVKRML 570
Db 770 ENHIDEIKGIDKLKKVPKNTWFPNEDPKIEKKIENIV---EKIDKKNIYKEI---DKL 823
Qy 571 LKXCIYGVNDINPISVEITMLSLWIN-TFIFGTPPLSF1-----EHHIKAGNALL 617
Db 824 LN-----EISKIENDKTSLEKLNINLSYSGSLGNLFLQIDEEKKAAEHTIKAMEA-- 875
Qy 618 GYTKDEFFDIVKKFESGFSFKGIKEIITILEDIYOKIKGIN-----DITKED 667
Db 876 -YIDD--LDNIIKKKSQE-----IEKEMNINMDIKMDIHKEMKALNISHDDYKIYHTTSKNH 928
Qy 668 IEK-----SKKIYKEVEESKIDONLRIIFSLIKLYSLSPDKSLNMEFSDIASVLSLIE 720
Db 929 EEKIDIRKNSUKIIQDFSEESYINDIKKELEKNVLESQNNNTDINQYLSKENIYNIIL- 987
Qy 721 NILNGKNTSSEDEKIEKIKLSYKFFHYGIEFPDIOEGFDIVIGNPPWEKTKFNETEF 780

Db 988 -----KLKIKKIIDKVK--EYTDIEIKNNKKNINAEIENSEKIITQIKENSS 1032

QY 781 FS-----KHIPNRYKGLKGEONIKIKEIILSKONHPLSIEYNEKNS 821

Db 1033 LKEQSKIKSTIDDNVSECINITNLTKYIIVNEKNINITYPKNAE-----EYNQ--NV 1084

QY 822 ITAINNIYKDFK-----CFTSGGDPNLFYFVTNFKLIKEKGNLTYLVPISATWNESS 875

Db 1085 SLNFNIEWADTKSYILNIKNKNGTNN-----TDYNIKELKEHKKKNVYKDEAGKNTQ 1139

QY 876 SRLRKHIPARYK-----LNYIYOFENKRF-----KDVHSSF----- 908

Db 1140 EIKKNKELFEKVEQVTVLKNKYAVELKNKFKDKYNYSEQIIEIKDAHNTFTSQADKS 1199

QY 909 -----KFAIFOLN-----IKESTSSFKAF-----MIQSSDNILKEIT 942

Db 1200 EKOMEIKNEQIRIEDEVAKNKSNKAILDIQLSVEPPKIKPLKTKLRTKSSDCLKE--T 1258

QY 943 RDLK-----DSKDAYKGI 956

Db 1259 KDIEFKISNLSIDTQETKLIENKNILNLEKLESKKNQKNIEQKELDEVNSKIKNI 1318

QY 957 ELNINQIKK-----LSPIQESIIEFKDNE--EPTLNK 987

Db 1319 ESNVQHKKNYEGIVEKINEITAKANKQIOESTQKLIPTIKNLISPPKANDLEGIDYTK 1378

QY 988 MFSKPSA--LGEYIDFPKGLD-----PSIKNRKSLKCEKNNKNI-----FLYSGANIHQ 1036

Db 1379 NLGKYNTEMNNIYEFIKSYDLITHYLETVSKEPITYEQIKNKRTAQWELLNINIKVWK 1438

QY 1037 FNSRFEDKAKESKLLWIDKEDLEKVLTK--DNQYQOTERVYRAIAS-----NTHER 1088

Db 1439 AKS--YLDIEANEPRIIVTHFPKNKLVNDKFTNEYSKVNKGFDNISNNVKKSTDEN 1497

QY 1089 TMISTLSP-----GNCYCVNSIYINDEK-----TPISLYKCLFI 1122

Db 1498 LLNLINQTKENYANIVSKYYSYKYEAENIFINIPKLANSLNIQIKSSGIDLPFNINI 1557

QY 1123 -----ISFNSFVDFLLR----- 1136

Db 1558 AILPYLDSKQDTLTFIPSEKTSYTYTKISDSYNTLLDLKRSQELQKQOQALNLIFE 1617

QY 1137 -----RFVDSNVLSKLYOCMPQOPEKEILSNFLYNLAKVTSLLTAKNDPENFKYLL 1190

Db 1618 NLLHDKVQATWELKDTLSL-----KNKKEQILNKVKLLHKSNELNKLSQNSQNYDTI- 1672

QY 1191 VLEYKFKDKVKYNKILKLDKDEPFKEKENEN--NFIASLYSLAKEDFTLLGDFKALKN 1249

Db 1673 -LESSYDKIK-----EKSNNYKEKENLGINFDVRAK-----BEQFNNDIKOIEKLEN 1720

QY 1250 KKKGEDIYSSLIKGYDNVLL-----NNKIFYHK 1277

Db 1721 NYKHSE-----KDNYNFSENNNLOS 1743

RESULT 12

F91286

hypothetical protein ECs5262 [imported] - Escherichia coli (strain O157:H7, substrain R1

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C:Accession: F91286

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene

A:Reference number: A99629; PMID:21156231; PMID:11258796

A:Accession: F91286

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1644 <HAV>

A:Cross-references: UNIPROT:Q8XC72; UNIPARC:UPI00000D04EE; GB:BA000007; PIDN:BA83685.1;

A:Experimental source: strain O157:H7, substrain R1MD 0509952

C:Genetics:

A:Gene: ECs5262

Query Match 5.7%; Score 371.5; DB 2; Length 1644;

Best Local Similarity 17.7%; Pred. No. 6.7e-07;

Matches 246; Conservative 231; Mismatches 47; Indels 443; Gaps 51;

QY 74 QKAGQIEGVSERVDILLPDKKASFNKILKEAKKNPIPIEDIL---IAEVKRTPE 130

Db 143 QEEGTLDPLAUSLTAQFPADTDKHGDSLKKKA--NGEYRSWQDLSTAVFTQNEPRF 200

QY 131 SFDKDKLKESEDQLYRYLNQYKHGILSKNGKVMRLYDKSKVLYGERRYIEFNFKIEE 190

Db 201 V-----LLGNRQLLLDRTK--WAQNRLRDFDEEILS 232

QY 191 KEEYQEQWFLFYILIRKERYLTKSNVIEVEKQIAKEKEI--IQTKLKEILVERPDD 248

Db 233 R---RETTDLKATALLHKDLSLPGSGAPYLDSDNDSHKHAFGVSEDLKVALRE---- 284

QY 249 IVFKIANNIYDFKLSKEITQHIL-----ASILEESIIFILRIFFIAY 293

Db 285 -----SIELGNEMAHYLIIDRGLANYTGNRAVDPELSRECLRYMYRLFLFY 332

QY 294 IEDNDIPKKILOENKLYRSSISFRY-----PFYDENTKKKLEYKKIITIFNLLDKGS 345

Db 333 IEARPELGYAPWTAQTYLQYSLETLDLEMIPLTSEEDRNGRYFHDLSNMLFKLVREG- 391

QY 346 DAIRPPVFGG-----LFSEDKVKYLNNEGLLSISEIE-----EILVKMLFFBEK 390

Db 392 -----YNGGVKMQSDLESGRDITIHSQ--FSVPRLESHLFDANNTILNRVVRNET 442

QY 391 -----NIKDEKFKYSLRDPKSPGELYETILEYDLRLADTTVHRI---- 430

Db 443 LQOIQAMSLSRPAKGRNRRGRISYRQLGINQLGAVYEAALLSYRGFFASEBDLYEVKAG 502

QY 431 -----IEDGVVLIKT-----EELNKKVKNKIATYKNGIYLTSRSLDRKSKGAYTTPD 479

Db 503 EEFNELEGTGVFVKDEISKYHDEKYEKDSLRHKGSTIYRMAGDRDKRSASYTPE 562

QY 480 DLTPMVISIEE-----QLTKYSPLDIKIIDNSCGSHFLISCLDYITEKVM 527

Db 563 VLTRSLVKYALKELFKQIDPISPHAKADAILNITVCEPAMGSAAFLEAINQLAEAYL 622

QY 528 YELDKFEDVKELDEEYRVIIIESEEDVDQDSISKELV--LKRMLLKXCIYGVDPISVE 586

Db 623 FH-----KQQAEGRI-----PODRTYQELQVRVMIADNNVFGVDLNPVAVE 665

QY 587 ITMLSLMINTFIPTPLSFIHHIKAGNALLG-----YTKDBFF----- 625

Db 666 LAEVSMLNALSIGDAFPWFQYQLHCGNSLVGARQVFNKSELITYKKADPSWLNSEPVE 725

QY 626 -----DIVKKCFESGFSFLFKRIKEI-----ITIL 650

Db 726 LAMNTPREETQIFHFLPDGGWMYVSDTKVQRYPDDFKALDSWEKFIKSPAGHEIADV 785

QY 651 EDIYQIKINGINDTKEDIE-----KSKKIYKEVEESKD 683

Db 786 QRISKEVEALWNTRYQQLKAERLKTADSYVPWPAENSEQTRSSLSKDETFSGRLDNSA 845

QY 684 IDNRILIFSL--IKLYSLSPKSLNMFSDIASVLSIENILGNKTSSEDEKIEK----- 737

Db 846 YOKLRWVMDYWCALWFWDIDKA--DELPDRGTWLFIEITLDDGIVITEKVTVAEHTTGD 903

QY 738 -----IRKLSSYY-----KPFHYGIFPPDI---QE 759

Db 904 LFAEGLLRRESSLFSVAGRLKTEVLFPHLPRLAIVDALRKQHRFFHWDLFFCDLFAERG 963

QY 760 GFDIVIGNPPWEKTKFNTEFFSKHPIPNY--RKLGIKEQNIKIKOILSKONHPLSIEYNE 817

Db 964 GFDMLGNPPWPKVQEWQAGVLDYEPFVLRLKSASKLATLRIDTF----- 1010

QY 818 EKNSIIANNIYKDFK-C-----FTSGGDPNLFYFVTNFKLIKEKGNL 862

Db 1011 --NQIPALEAAWRSEYEGCEGQNFNAQNYPVLRGVQTNLYKCFLPQAWRLGAOKGVA 1068

Query Match		5.5%;	Score 359;	DB 2;	Length 2166;
Best Local Similarity		21.4%;	Pred. No. 2.8e-06;		
Matches		325;	Conservative	245;	Mismatches 473; Indels 474; Gaps 87;
QY	2	KTNDIVKTNPNISLYKQSKDPFKENI-----NKLKDPFILKNKLFSDDDNST	52		
Db	769	RQTEIHK--DKSVFMLEDLNNKFFDDKNFVSKIEBCDYKLKDFKIESDILNFKSDLN	826		
QY	53	ENIESLLXYIEELNYSVEQKAGIEGVSERV--DILLFENDKDKASFNNKLK-----	105		
Db	827	EF-IESKQIV---SNIKSDNQ--QIDDFLDRIKDIIL---NRKD--SINNEVSKLSD	875		
QY	106	BAKQNEPIEDILIAEVKRPFPFADAKLKE--SEDOLVRY---LNOYQKHVG	157		
Db	876	WQSKLNEITVKLENLSSKVDLDLIDSEVTYKIKELFESIESLYLEKIDEPNQA	935		
QY	158	ILSNGKWRVLYDKSVLYGKERYIEFNFXK-----IEEKERY-KQEQFVFLYIIRKER	211		
Db	936	IYSDE---LQDIDMNHFNKETRELEENLSKKFAAVLANSEEFVKEVD-----SLQDKR	986		
QY	212	-----YLKTSNV--LEVEKQIAKEKEII-----QXTLKEIILYE---	243		
Db	987	TDIASQANIDITLDSIANVKFNDINKEINGKYNEISYRGENISSKLENEIMHEIEN	1046		
QY	244	---RPDDSI-----VFKIAXNYDKPEKPLSGKEIT-----	270		
Db	1047	LSRLTDRDLSKGMENDLQKLKSFVSKYQVEK-FELKVKDLTDGGEAKINKLVKEI	1105		
QY	271	QHILASILEESIIFILRIFPIAYIEDNDI-----FKILOENKLYRSSISPRYFFYDEN	324		
Db	1106	EQYKSRLEEAIDYRTI-----DNDIMQAKERFGEITNELK--NNIESKSEFLNDL	1155		
QY	325	TKKLEYKKIITIFNLLDKGSDAIFPVNGGLFBSDKVYLNNEGLLSISIEBILVKM	384		
Db	1156	YKER-----FKLIESNFEE-RYSTF-----LIESEG--AISKIRDEIYKT	1192		
QY	385	LPFEEKNIKDEKVFVYSRLDKPSFGELYE---TLLEYDLRIAD-----	424		
Db	1193	LTSNDENLQ---IKISEMD-QNF-EILEQRKDIIEFEKELQDKIKDCYGFINSQFGEI	1246		
QY	425	-----TTVHRIIEGCVYLIRTEEELEENKVNKIATYLNKNTYLTSRSLD	468		
Db	1247	KAGVERNINKHFDVCIKVTNLTIDDI--VKYENEI-HKRIDSLKS-IESTFDSIEKNLN	1302		
QY	469	RKSGAYTTPDLDLTPWVSSI-----EEOIKTKSPLD--IKIIDNSCGSGHFLISCL	519		
Db	1303	DKVSGCI--DKIANDFNLYKYLEERCNEGOLNLENKIDNKIKAIN-----LALSQY	1353		
QY	520	DYLTEKVVYELDKFEDVKELDEYRVIIIEESEYDVQDSISKELV--LKRMLLKKCIYG	577		
Db	1354	DGLERKY---ADMYDEFSERLNSYIATLSEEPK-----SSNKEWIFELESQKNLKNLE	1404		
QY	578	VDINPISVEITMLSLWINTFIFGTPLSFTEHHIKAGNALLGYTKOEFDIVKKKFSFGFS	637		
Db	1405	SDLN--NVEKDVIRLKEEY-----HNVSSHLLKL---EEDFFKDLKIRGEE---	1446		
QY	638	LPFKRIKEIITILEDIYQKIKINDTTKEDIKSKKIYKEEESKIDINLRIIFSLLIK-	696		
Db	1447	LKYSLENFIASYNKIQNL-----EYDLKSNLENKTELIOQSFRLD	1486		
QY	697	-----YSLSPDKSLNMFSDIASVISIENILGNKTSSEDEKIEKIRKLSSYY	745		
Db	1487	IEQKKKDDKENFVLDTKBFSSKKDKMQSEIALMET---NITGKVD-EFVDFVNNKQSII	1542		
QY	746	KFFHYGIEFPDIOEGPDIIVIGNPPWEKTKFNTEFPFSKHIPNRYKLGIGK--EQNIIKOEI	803		
Db	1543	DSWFENIK-DDVKD-----WQEKSYSTIE---KEI-NLAELGKISFENDIFNVKI	1587		
QY	804	----LSKDNHPLSIE--YNEEKSIITAINNIYKDFPKCFTSGGDPNL-----FRYFTFNLK	854		
Db	1588	GLESFMDGFEIKAEEIFSNLQNEAKKIEQSVHLDFK--NIGESLNLKVLDEKFDVDFKLE	1645		
QY	855	LIKEGNL---TYLVPSAIWNNESSRILRKHIFARYKLVNIYQFE-----NKKRF	901		

Search completed: January 24, 2006, 19:59:13
Job time : 55.4458 secs

Db	1646	KIDEKVNKKTEDILLQAQVVKFLTQCKOLEDKIF--ELNQKLEHEFTTILSSNLDKVRREM	1702		
QY	902	KDVHSSFKFPAI-FQLSNIKESTSSSFKAQFM-----IQSS-----DNILKEITRDLKSKD	950		
Db	1703	VDVISSDKESFEGQTEILNKIISBFSEKISLYRNNIETSIENEVNFSKSIKSDGLLGLD	1762		
QY	951	DAYGIELNINQIKKL-SPIQESI-----IEFK-----DNEEFTLINKM-----F	989		
Db	1763	ELKXSLRHSTSEIETIKSGLOEQDKDFEVEFKONHKELLKEVDNNILESKILMCDVOF	1822		
QY	990	SKF-SALGEGVIDFKKGL-----DPSIKNRKSLLKECNCNNKILFLYSGAN	1033		
Db	1823	NKFISEIKDNLVVEYKSDLRAEFEDSYDKINFIQENIENFKKLDSELEKNNSIFL----	1877		
QY	1034	IHQFNRSFFEDKADAKESSKLLWIDKEDLEKVL-----TKDNOYQTERVYFVYRAIASNTNER	1088		
Db	1878	-----EAYSLLDKLEKLW--ETLKNEIGLAQYKNNFENVNKEFY-----NIQKE	1920		
QY	1089	TWISTLSPGNCYCNSIYINDEKTPISLYKKLFIISIENSP--VPDF-----LLR	1136		
Db	1921	TL-----GIIEIFNELKLEQES-----IKSIKNDFNRFPEFYSSFSDSRYSKSLIB	1964		
QY	1137	RFVDSNVLSCLYOCMPQPEEKEILSNPLYNLAKNTSLLIAKNDPENFKYLLYLEVFK	1196		
Db	1965	SYDEMOIYKAKIKEI--ADEQRTILDN--YERISNKESIL-----	2000		
QY	1197	FDKEKVNKILKLDKEDFEKKEKENENNFIASLYSLAKEDFITLLGDFKALKNKKKGBDY	1256		
Db	2001	-----KSTIESVDKNFDLINEVEKRFN-----NLSKE-----SAKIQDNLMKMEV	2041		
QY	1257	ISSLI--KGYDNYLLNN	1271		
Db	2042	VSSLLLNKGLSEEVLIN	2058		

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:27:53 ; Search time 274.4 Seconds
(without alignments)
3283.383 Million cell updates/sec

Title: US-10-688-058-16
Perfect score: 6538
Sequence: 1 MKTNDIVKTNPNISLYKQL.....SSLIKGYDNYLLNNKIFVHK 1277

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	6534	99.9	1277	2	050698 BORBU
2	5599	85.6	1278	2	050667 BORBU
3	5472	83.7	1194	2	05XYT2 BORGA
4	3119	47.7	806	2	05XYG5 BORGA
5	2648	40.5	635	2	05XYI7 BORGA
6	1781	27.2	1250	2	09PPL7 CAMJF
7	1764.5	27.0	1250	2	05HV91 CAMJR
8	1648.5	25.2	422	2	050806 BORBU
9	671	10.3	149	2	05XYI6 BORGA
10	605.5	9.3	1318	2	04HA54_9DEIO
11	604	9.2	1442	2	07NGZ7 GLOVI
12	558.5	8.5	706	2	052000 HALSA
13	546.5	8.4	662	2	09HI06 HALSA
14	546.5	8.4	1610	2	05FOU8 GLUOX
15	526.5	8.1	1181	1	Y3402_METJA
16	525.5	8.0	1186	2	04HL25 CAMLA
17	522.5	8.0	1257	2	05HXC7 CAMJR
18	510.5	7.8	1255	2	04HJ16 CAMLA
19	488	7.5	878	2	04HEH3 CAMCO
20	473	7.2	134	2	050807 BORBU
21	455.5	7.0	1252	2	09ZJA4 HELPJ
22	446	6.8	867	2	09PJ80 CAMJF
23	445	6.8	1203	2	08E4S7 STRA3
24	435	6.7	1532	2	08NSD8 CORGL
25	433	6.6	1532	2	06M742 CORGL
26	419.5	6.4	1279	2	026046 HELPY
27	405.5	6.2	5729	2	081622 PLAF7
28	398	6.1	2664	2	07RELO PLAYO
29	395.5	6.0	2723	2	07RQB6 PLAYO
30	395	6.0	2965	2	08WRS5 PLAF6
31	394	6.0	2771	2	026216_9APIC

32	392	6.0	2957	2	08WRS6 PLAF6	Q8WRS6 plasmodium
33	392	6.0	5779	2	08IBS0 PLAF7	Q8IBS0 plasmodium
34	391	6.0	2770	2	07YUE9 PLAF6	Q7YUE9 plasmodium
35	390	6.0	2976	2	08WRS4 PLAF6	Q8WRS4 plasmodium
36	389.5	6.0	1513	2	04YR36 PLABE	Q4YR36 plasmodium
37	388	5.9	2969	2	07K973 PLAF6	Q7K973 plasmodium
38	386.5	5.9	2977	2	08WE99 PLAF6	Q8WE99 plasmodium
39	385.5	5.9	2749	2	07REY3 PLAYO	Q7REY3 plasmodium
40	385.5	5.9	2752	2	09BJY0 PLAYO	Q9BJY0 plasmodium
41	382	5.8	1009	2	08KRW6 BACPU	Q8KRW6 bacillus pu
42	379.5	5.8	1414	2	063WK7 BURPS	Q63WK7 burkholderi
43	379	5.8	2881	2	06YA78 PLABE	Q6YA78 plasmodium
44	378	5.8	1524	2	08FSZ4 COREF	Q8FSZ4 corynebacte
45	378	5.8	1993	2	08ILC6 PLAF7	Q8ILC6 plasmodium

ALIGNMENTS

RESULT 1
O50698 BORBU PRELIMINARY; PRT; 1277 AA.
ID O50698; AC O50698;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hypothetical protein BBE02.
GN OrderedLocNames=BBE02;
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp25.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=139;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Fraser C.M., Casjens S., Huang M.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
RA Gwin M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Weidman J.F., Utterback T.R., Wattney L., McDonald L.A.,
RA Artach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi."
RL Nature 390:580-586 (1997).
DR EMBL; AE000785; AAC66031.1; -; Genomic_DNA.
DR PIR; E70224; E70224.
DR HSSP; Q57997; 1MJH.
DR TIGR; BBE02; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0006308; F:DNA methylation; IEA.
DR InterPro; IPR007409; DUF450.
DR InterPro; IPR002296; N12N6 mtfase.
DR InterPro; IPR002052; N6 Mtfase.
DR Pfam; PF04313; HSDR_N; 1.
DR PRINTS; PRO0507; N12N6MTFRASE.
DR PROSITE; PS00092; N6 MTFASE; UNKNOWN_1.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 1277 AA; 150972 MW; 0B41DF4EDB5859F4 CRC64;

Query Match 99.9%; Score 6534; DB 2; Length 1277;
Best Local Similarity 100.0%; Pred. No. 1.5e-243;
Matches 1277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKTNDIVKTNPNISLYKQSKDFIKENINKLKDFFILIKNLFSDIDNSTEAMIESLL 60
DB 1 MKTNDIVKTNPNISLYKQSKDFIKENINKLKDFFILIKNLFSDIDNSTEAMIESLL 60
QY 61 KYIFBELNYSVEQKAGQIEGVSVDILLFENDKDKAFNNKKEAKQNPPIEDIL 120

Db 192 SKIKEBYKEQWFIPLSYLIRKERYLTSNIIISVEKEQISKEKEIIQKTLREILYERP 251
Qy 246 DDSIVFKIAKNYDKPKLKGKQTHILASILEESIIFILRIFFIAYIEDNDIPKILQ 305
Db 252 DDSIVFKIAKNYDKPKLKGKQTHILASILEESIIFILRIFFIAYIEDNDIPKILQ 311
Qy 306 ENKLYRSISPRYPFYDENTKGLYKIIITFNLDKGSDAIKPPVNGGLPSEDVKY 365
Db 312 ENKLYRSISPRYPFYDENTKGLYKIIITFNLDKGSDAIKPPVNGGLPSEDVKY 371
Qy 366 LNNEGLLSISIEEILVKMLPFEENKIDKFKVYKSRDPSFGELYETLLEYDLRIADT 425
Db 372 LNNEGLLSISIEEILVKMLPFEENKIDKFKVYKSRDPSFGELYETLLEYDLRIADT 431
Qy 426 TVHRIEDGVYLIRTEBELENKKNKIATYKGNIIYLSRDLKSKGAYTTPDLDPM 485
Db 432 TVHRIEDGVYLIRTEBELENKKNKIATYKGNIIYLSRDLKSKGAYTTPDLDPM 491
Qy 486 VISSTEEOLKTSPLDIKIIDNSCGSHPLISCLDYLTEKWYELDKEDVKELDEYR 545
Db 492 VISSTEEOLKTSPLDIKIIDNSCGSHPLISCLDYLTEKWYELDKEDVKELDEYR 551
Qy 546 VIESEYVDQDSISKELVKMLKXCIYGVNDINPISVEITMLSLWINTFIQTPLSF 605
Db 552 IILKESEYVDQDSISKELVKMLKXCIYGVNDINPISVEITMLSLWINTFIQTPLSF 611
Qy 606 IEHHKAGNALLGYTKDBFFDIIVKKGPSGSLFKKRIKEIITILEDYQKIKGINDTTK 665
Db 612 IEHHKAGNALLGYTKDBFFDIIVKKGPSGSLFKKRIKEIITILEDYQKIKGINDTTK 671
Qy 666 EDIEKSKIKYEEESKIDNRIIIFSLIKLISLSDKSLNMFSDIASVLSLIENILGN 725
Db 672 EDIEKSKIKYEEESKIDNRIIIFSLIKLISLSDKSLNMFSDIASVLSLIENILGN 731
Qy 726 KTSSEDEKIKIRKLSVYFFHYGIEFPDIQEGFDIVGNPPWEKTFNETEPFSKHI 785
Db 732 KTSSEDEKIKIRKLSVYFFHYGIEFPDIQEGFDIVGNPPWEKTFNETEPFSKHI 791
Qy 786 PNRYKLGIKEQNIIOBILSKDNHPLSIEYNEEKNSIIAINNIYKDFKCTSGGDPNLF 845
Db 792 PSYRKLSIKEQNKIQEILGRDNHPLSIEYNEEKNSMGITNNLYKSDPKNFASGGDPNLF 851
Qy 846 RYFVTFNKLKEKGNLYLPSALWNESSRILRKHFARYKLYNIYQEPENKPKFQVH 905
Db 852 RYFVAFNLKIKENGNTLYLPSALWNESSRILRKHFARYKLYNIYQEPENKPKFQVH 911
Qy 906 SSFKAIFQLSNIKESTSEKAKFMIOQSDNLIKETIQLDKSDKDAYKGIELNINQIK 965
Db 912 TLFKFAIFQISNTKPTSNFKAIFMIOQSDNLIKETIQLDKSDKDAYKGIELNINQIK 971
Qy 966 LSPQIESIIEFKDNEEFTLINKWPSKFGALGEGYIDFKKGLDPSIKNRKSLIKECNNKL 1025
Db 972 LSPQIESIIEFKDNEEFTLINKWPSKFGALGEGYIDFKKGLDPSIKNRKSLIKECNNKL 1029
Qy 1026 IFLYGANTHQNPSRFFEDKAKESKLLWIDKEDKLVTKDNQYQTERVYFAIASNT 1085
Db 1030 IFLYGANTHQNPSRFFEDKAKESKLLWIDKEDKLVTKDNQYQTERVYFAIASNT 1089
Qy 1086 NERTWISTLSPGNCYCVNSIYINDEKTPISLYKLPISISFNSFYVDLLERF-VDSNVL 1144
Db 1090 DIRTWISTLSPGNCYCVNSIYINDEKTPISLYKLPISISFNSFYVDLLERF-VDSNVL 1149
Qy 1145 KSCLQCPMPQPEEKEILSNPLYLAKNTSLIAKNPDPENPKYLLYLEYFKFDEKVNK 1204
Db 1150 KSCLQCPMPQPEEKEILSNPLYLAKNTSLIAKNPDPENPKYLLYLEYFKFDEKVNK 1209
Qy 1205 ILKLDKEDFPFKEKENNFIITASIYSLAKEDFITLLGDFKALKNNKKGEDYISSIKGY 1264
Db 1210 ILKLDKEDFPFKEKENNFIITASIYSLAKEDFITLLGDFKALKNNKKGEDYISSIKGY 1268
Qy 1265 DNYLNNKI 1273
Db 1269 ENYLKTNKL 1277

RESULT 3
Q5XYT2 BORGA PRELIMINARY; PRT; 1194 AA.
ID Q5XYT2 BORGA PRELIMINARY; PRT; 1194 AA.
AC Q5XYT2;
DT 25-OCT-2004 (Tremblrel. 28, Created)
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=BGP199;
OS Borrelia garinii PBI.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_taxid=290434;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBI;
RA Gloeckner G., Schilhabel M., Lehmann R., Platzer M.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY722925; AAU86050.1; -; Genomic_DNA.
DR GO; GO:0003777; F:DNA binding; IEA.
DR GO; GO:0008170; P:N-methyltransferase activity; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR InterPro; IPR007409; DUF450.
DR InterPro; IPR002296; N12N6 mtfase.
DR InterPro; IPR002052; N6 Mtfase.
DR Pfam; PF04313; HSDR N; 1.
DR PRINTS; PR00507; N12N6MTFASR.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 1194 AA; 141642 MW; 51DC343D205F2B41 CRC64;

Query Match 83.7%; Score 5472; DB 2; Length 1194;
Best Local Similarity 88.5%; Pred. No. 1.1e-202;
Matches 1056; Conservative 74; Mismatches 63; Indels 0; Gaps 0;
Qy 1 MKTNDIVKTNPNISLYKQLSKDFIKENINKLKDFFILIRKKNLFSIDNTEANIESLL 60
Db 1 MKTNNIVKTNPNISLYKQLSKDFIKENINKLKDFFILIRKKNLFSIDNTEANIESLL 60
Qy 61 KYIFEELNYSVEQKAGQIEGVESRVDILLFENDKDKASFNKLKEAKNNEPIEDIL 120
Db 61 RSIPEELNYSVEQKAGQIEGVESRVDILLFENDKDKASFNKLKEAKNNEPIEDIL 120
Qy 121 IIAEVRPTFSFADAKDLKESDQLYRYLNOYQHYGILSNQKWLRYDKSKVLYGKRY 180
Db 121 IIAEVRPTFSFADAKDLKESDQLYRYLNOYQHYGILSNQKWLRYDKSKVLYGKRY 180
Qy 121 IIAEVRPTFSFADAKDLKESDQLYRYLNOYQHYGILSNQKWLRYDKSKVLYGKRY 180
Db 121 IIAEVRPTFSFADAKDLKESDQLYRYLNOYQHYGILSNQKWLRYDKSKVLYGKRY 180
Qy 181 IEFNFKTEEEKEEYKQWFWFIYILIRKERYLTSNIIISVEKEQISKEKEIIQKTLREI 240
Db 181 IEFNFKTEEEKEEYKQWFWFIYILIRKERYLTSNIIISVEKEQISKEKEIIQKTLREI 240
Qy 241 LYEPDDDSIVFKIAKNYDKPKLKGKQTHILASILEESIIFILRIFFIAYIEDNDIF 300
Db 241 LYEPDDDSIVFKIAKNYDKPKLKGKQTHILASILEESIIFILRIFFIAYIEDNDIF 300
Qy 301 KKLQENKLYKSSISFRYPFYDENTKGLYKIIITFNLDKGSDAIKPPVNGGLPSE 360
Db 301 KKLQENKLYKSSISFRYPFYDENTKGLYKIIITFNLDKGSDAIKPPVNGGLPSE 360
Qy 361 DKVKYLNNEGLLSISIEEILVKMLPFEENKIDKFKVYKSRDPSFGELYETLLEYDL 420
Db 361 DKVKYLNNEGLLSISIEEILVKMLPFEENKIDKFKVYKSRDPSFGELYETLLEYDL 420
Qy 421 RIADTTVHRIEDGVYLIRTEBELENKKNKIATYKGNIIYLSRDLKSKGAYTTPDD 480

Db 421 RIADTTVHRIVEDEVYLRTEBELENKKNVATYKGNILYTSRLNRKSGAYTTPDD 480
QY 481 LTDWFVISSIEQLTKTSPLDIKIIDNSCGSHFLISCLDYLTETKVVWYELDFEDVKKEL 540
Db 481 LTDWFVISSIEQLTKTSPLDIKIIDNSCGSHFLISCLDYLTETKVVWYELDFEDVKKEL 540
QY 541 DDEYRVIIIESEYDQDSISKELVLRMLKXKCYGVDPINPISVEITWLSWINTFTFG 600
Db 541 DDEYRVIIIESEYDQDSISKELVLRMLKXKCYGVDPINPISVEITWLSWINTFTFG 600
QY 601 TPLSFIEHHIKAGNALLGYTDEFFDI VVKKPFESGFLFKRIKRIITILEDIYQIKIGI 660
Db 601 TPLSFIEHHIKAGNALLGYTDEFFDI VVKKPFESGFLFKRIKRIITILEDIYQIKIGI 660
QY 661 NDTTKEDIEKSKIKIYEVEESKDIDNLRIFSLIKYLSLSPKSLNRPESDIASVISLIE 720
Db 661 NDTTKEDIEKSKIKIYEVEESKDIDNLRIFSLIKYLSLSPKSLNRPESDIATVISLIE 720
QY 721 NILGNKTSSEDEKIEKIRKLSYKFFHYGIEFPDIOEGFDIVIGNPPWEKTKNETEF 780
Db 721 NILGNKTSSEDEKIEKIRKLSYKFFHYGIEFPDIOEGFDIVIGNPPWEKTKNETEF 780
QY 781 FSKHIPNYRKLGIKEQNIIOBILSKONHPLSIEYNEEKNSIIANNIYKFDKFCGTSG 840
Db 781 FSKHIPNYRKLGIKEQNIIOBILSKONHPLSIEYNEEKNSIIANNIYKFDKFCGTSG 840
QY 841 DPNLFRYFTVNLKIKKGNLTYPVPSAINWESSRILRKHI PARYKLNIIYQFENKCR 900
Db 841 DPNLFRYFTVNLKIKKGNLTYPVPSAINWESSRILRKHI PARYKLNIIYQFENKCR 900
QY 901 FKDVHSSFFALFQLSNIKESTSSFKAKPMIOSSDNILKEITRDLDKSDKDAYKGIELNI 960
Db 901 FKDVHSSFFALFQLSNIKESTSSFKAKPMIOSSDNILKEITRDLDKSDKDAYKGIELNI 960
QY 961 NOIKKLSPTQESIIIEFKDNEETFLINKMPSKFSALGEGYIDFKKGLDPSIKNRKSLKEC 1020
Db 961 NOIKKLSPTQESIIIEFKDNEETFLINKMPSKFSALGEGYIDFKKGLDPSIKNRKSLKEC 1020
QY 1021 NNKLIIFLYSGANIHOFSRPFEDKDAKESKLLMIDKEDLEKVLTKDQYOTERVFYRA 1080
Db 1021 NNKLIIFLYSGANIHOFSRPFEDKDAKESKLLMIDKEDLEKVLTKDQYOTERVFYRA 1080
QY 1081 IASNTWERTWISTLSPGNCVSNIIYINDEKTPISLYKLLFIISFNSFVDFLLRRFVD 1140
Db 1081 IASNTWERTWISTLSPGNCVSNIIYINDEKTPISLYKLLFIISFNSFVDFLLRRFVD 1140
QY 1141 SNVLKSLYQCPMPQPEEKEILSNPLNLAKNSTSLIAKNDPENPKYLLYLE 1193
Db 1141 SNVLKSLYQCPMPQPEEKEILSNPLNLAKNSTSLIAKNDPENPKYLLYLE 1193

RESULT 4
Q5XYG5 BORGAC PRELIMINARY; PRT; 806 AA.
AC Q5XYG5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein (Fragment).
GN ORFNames=BGP316;
OS Borrelia garinii PBI.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_Taxid=290434;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBI;
RA Gloeckner G., Schilhabel M., Lehmann R., Pradella S.,
RA Schulte-Spechtel U., Schilhabel M., Wilske B., Sühnel J., Platzer M.;
RL "Comparative analysis of the Borrelia garinii genome."
RL Nucleic Acids Res. 32:6038-6046(2004).
RN [2]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBI;
RA Gloeckner G., Schilhabel M., Lehmann R., Platzer M.;
RL Submitted (SRP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY722945; AAU86167.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; P:N-methyltransferase activity; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR InterPro; IPR002296; N12N6_mtfase.
DR InterPro; IPR002052; N6_Mtase.
DR PRINTS; PR00507; N12N6MTFRASE.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
KW Hypothetical protein.
FT NON TER 806 806
SQ SEQUENCE 806 AA; 94261 MW; 2590359906792188 CRC64;
Query Match 47.7%; Score 3119; DB 2; Length 806;
Best Local Similarity 73.4%; Pred. No. 2.7e-112;
Matches 592; Conservative 104; Mismatches 108; Indels 2; Gaps 2;
QY 371 LLSIEIEEIIIVKMLFPEEKNI KDKFVKYSRLDPKSFGEYVETLLEVDLRADTTVHRI 430
Db 1 MLSISEEIIILIKLFLSEKNIKDKKFKVEYSLDKPKIFGEIYENLKKYDLRIADTTIHKI 60
QY 431 IEDGVYLRTEBELENKKNVATYKGNILYTSRLNRKSGAYTTPDDLTDFWVSSI 490
Db 61 IEKGTICYHTDELDKKNKKNVATYKGNILYTSRLNRKSGAYTTPEDLADFWAASAV 120
QY 491 BEQKTYGSPLDIKIIDNSCGSHFLISCLDYLTETKVVWYELDFEDVKKELDEEYRVIEE 550
Db 121 BEQKTKFPPLDIKIIDNSCGSNLLISCLDYLTETKVVYSLDKFEDLKKLDMEYRNVIDE 180
QY 551 SEEDVDVDSISKELVLRMLKXKCYGVDPINPISVEITWLSWINTFTFGTPLSFLIEHI 610
Db 181 AKKNIQDSVRKIKRILRKICIVGVDFDQISVQIAMLRLWIGVIFVSPSPSLTNRI 240
QY 611 KAGNALLGYTDEFFDI VVKKPFESGFLFKRIKRIITILEDIYQIKIGINDTTKEDIEK 670
Db 241 KVGNALLGYTDEFFDI VVKKPFESGFLFKRIKRIITILEDIYQIKIGINDTTKEDIEK 670
QY 671 SKKIYKEYESKODINLRIFSLIKYLSLSPKSLNRPESDIASVISIENILGNKTSSE 730
Db 301 SKKIYKEYESKODINLRIFSLIKYLSLSPKSLNRPESDIASVISIENILGNKTSSE 730
QY 731 DKEKIEKIRKLSYKFFHYGIEFPDIOEGFDIVIGNPPWEKTKNETEFESKHIPNYRK 790
Db 361 DKEKIEKIRKLSYKFFHYGIEFPDIOEGFDIVIGNPPWEKTKNETEFESKHIPNYRK 790
QY 791 LGIKEQNIIOBILSKONHPLSIEYNEEKNSIIANNIYKFDKFCGTSGDPLNRYFTV 850
Db 421 LSIKEQNIIOBILSKONHPLSIEYNEEKNSIIANNIYKFDKFCGTSGDPLNRYFTV 850
QY 851 FNLKLIKKEGNIITVLPVPSAINWESSRILRKHI PARYKLNIIYQFENKCRFKDVHSSPKF 910
Db 481 FNLKLVKPGGNLTYPVPSVLCCEKSKALRQHIFSNYKLNIIYQFENKCRFSVANSSEKF 540
QY 911 AIFQLSNIKESTSSFKAKPMIOSSDNILKEITRDLDKSDKDAYKGIELNINQIKLSPTQ 970
Db 541 AIFQTSNTKEPTTFRVRFVYQSCNIMKEITSDUKEGNNAYKGIELDIAQIKLSPTQ 600
QY 971 ESIIIEFKDNEETFLINKMPSKFSALGEGYIDFKKGLDPSIKNRKSLKECNNKLIIFY 1020
Db 601 ESIIIEFRDI VVKLIDKMFSSPSFSEDYMNFI GLNTADMNKLKFLKCEKDENCIFLY 660
QY 1030 SGANIHOFSRPFEDKDAKESKLLMIDKEDLEKVLTKDQYOTERVFYRAIASNTWERT 1089
Db 661 CGANIHOFSRPFEDKDAKESKLLMIDKEDLEKVLTKDQYOTERVFYRAIASNTWERT 1089
QY 1090 MISTLSPGNCVSNIIYINDEKTPISLYKLLFIISFNSFVDFLLRRFVD -SNVLKSL 1148
Db 721 MISTLSPGNCVSNIIYINDEKTPISLYKLLFIISFNSFVDFLLRRFVD -SNVLKSL 1148
QY 1149 YQCPMPQPEEKEILSNPLNLAKNSTSLIAKNDPENPKYLLYLE 1194

Db 781 YOCFMMQPEKKNILANPLTLTKNT 806

RESULT 5

Q5XY17_BORGA
ID Q5XY17_BORGA PRELIMINARY; PRT; 635 AA.

AC Q5XY17;
DB 25-OCT-2004 (TREMELrel. 28, Created)
DT 25-OCT-2004 (TREMELrel. 28, Last sequence update)
DE Hypothetical protein.
GN ORFNames=BG294;
OS Borrelia garinii PBI.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=290434;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBI;
RA Gloeckner G., Lehmann R., Romualdi A., Pradella S.,
RA Schulte-Spechtel U., Schilhabel M., Wilske B., Suhnel J., Platzer M.;
RT "Comparative analysis of the Borrelia garinii genome.";
RN Nucleic Acids Res. 32:6038-6046(2004).
[2]
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NCTC 11168;
RA MEDLINE=20150912; PubMed=10688204; DOI=10.1038/35001088;
RX Parkhill J., Wren B.W., Mungall K.L., Ketley J.M., Churcher C.M.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajararam M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
DR EMBL; AL139076; CAB72964.1; -; Genomic_DNA.
DR F1R; E81339; E81339.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR InterPro; IPR007409; DUF450.
DR InterPro; IPR002296; N12N6_mtfase.
DR InterPro; IPR002052; N6_Mtase.
DR Pfam; PF04313; HSDR_N; 1.
DR PRINTS; PR00507; N12N6MTFRASE.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 1250 AA; 147698 MW; 0A2CD9C5A464CDEA CRC64;

Query Match 40.5%; Score 2648; DB 2; Length 635;
Best Local Similarity 81.0%; Pred. No. 2.8e-94;
Matches 516; Conservative 44; Mismatches 75; Indels 2; Gaps 1;

Qy 569 MLLKXCIYGVNDINPISVEITMLSLWINTIFGTPLSFIEHHIKAGNALLGYTKDEFDVI 628
Db 1 MLLKXCIYGVNDINPISVEITMLSLWINTIFGTPLSFIEHHIKAGNALLGYTKDEFDIT 60
Qy 629 KKFESGSLFKKRIKEITILEDYQKTINGINDTKGDIKSKKIKEYEESKDINLR 688
Db 61 KKFESGSLFKKRIKEITILEDYQKTINGINDTKGDIKSKKIKEYEESKDINLR 120
Qy 689 IIFSLKLSLSPDKSLNMFSDIASVISLIENILGNKTSSEDEKIKRLSYKFF 748
Db 121 IIFSLKLSLSPDKSLNMFSDIASVISLIENILGNKTSSEDEKIKRLSYKFF 180
Qy 749 HYGIETPDQEGFDIVGNPPPEKTFNTEFFSKHINPYRKLGIKEQMIKOEILSKDN 808
Db 181 HYGIETPDQEGFDIVGNPPPEKTFNTEFFSKHINPYRKLGIKEQMIKOEILSKDN 240
Qy 809 HPLSIEYNEKNSIIANNIYKDFKFTSGGDPNLFYVTFNKLKEKGNLTLYLPS 868
Db 241 HPLSIEYNEKNSIIANNIYKDFKFTSGGDPNLFYVTFNKLKEKGNLTLYLPS 300
Qy 869 AIWNESSRLRGHIFARYKLVYIYFENKPKFVHVSFPFAIPOLSNIEKTSFPKAK 928
Db 301 AIWNESSRLRGHIFARYKLVYIYFENKPKFVHVSFPFAIPOLSNIEKTSFPKAK 360
Qy 929 FMIQSDNLIKETTRDLKQSDAYGIELNIQKLSPIQESIIEFKDNEEFTLNKM 988
Db 361 FMIQSDNLIKETTRDLKQSDAYGIELNIQKLSPIQESIIEFKDNEEFTLNKM 420
Qy 989 FSKFSALGEGYIDFKGLDPSINRKLKCECNKNLIFLYSGANIHFNSRFPEDKDAK 1048

Db 421 FTQFSILSEYIDFIFGL--NLTKNKKALLKESNNKNFIFLYCGANIHFNSRFPKDKSVK 478

Qy 1049 ESSKLLWIDKEDLEKVLTKQNYQOTFRVYRAIASNTINERTMTSLSPGNCYCYNSTIYN 1108

Db 479 ESSKSPWIDKEDLKKVLTEDNKKLNESIIYRKIARTDITMTSLCPKNSCYCIESLCIN 538

Qy 1109 DEKTPISLYKKLFIISIFNSVFDFLLRFRFVDSNVLSKSLYQCMPQOPEKEILSNPLYL 1168

Db 539 YEKTAISIYKKLFIISIFNSLAFDFLLRFRFVDSNVLSKSLYQCMPQOPEEKDILANSLYL 598

Qy 1169 NLAKNTSLLIAKNDPENFKYLLLYEYKFKDEKYNKI 1205

Db 599 TLTKNTSLLIAKNDPENFKYLLLYEYKFKDEKYNKI 635

RESULT 6

Q9PPL7_CAMJF
ID Q9PPL7_CAMJF PRELIMINARY; PRT; 1250 AA.

AC Q9PPL7;
DB 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DE Possible restriction /modification enzyme.
GN OrderedLocusNames=Cj0690c;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NCTC 11168;
RA MEDLINE=20150912; PubMed=10688204; DOI=10.1038/35001088;
RX Parkhill J., Wren B.W., Mungall K.L., Ketley J.M., Churcher C.M.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajararam M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
DR EMBL; AL139076; CAB72964.1; -; Genomic_DNA.
DR F1R; E81339; E81339.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR InterPro; IPR007409; DUF450.
DR InterPro; IPR002296; N12N6_mtfase.
DR InterPro; IPR002052; N6_Mtase.
DR Pfam; PF04313; HSDR_N; 1.
DR PRINTS; PR00507; N12N6MTFRASE.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 1250 AA; 147698 MW; 0A2CD9C5A464CDEA CRC64;

Query Match 27.2%; Score 1781; DB 2; Length 1250;
Best Local Similarity 34.7%; Pred. No. 1.4e-60;
Matches 457; Conservative 244; Mismatches 479; Indels 138; Gaps 40;

Qy 5 DIVKTNPNISLYKQLSKDFIKKENINKLKDFF-----ILIKNKLFSIDDNSTEAN 55
Db 6 ELKKYLNNNLFTSYLSLEIDFPKIFYNFANQKAFRDLTKITKYDKNKFQKNEHQHFE-- 63
Qy 56 IESLLKYIPEELNYSVEQKAGQIEGVESRVDILLFENDKDKAFNNKLEAKKNNEPIP 115
Db .64 -DEFTSKVLEILGWCFCVQDEKIQGLKEKDFDLFLSNDKLSKYENLDKETKSS----- 118
Qy 116 IEDILITAEVKRPTSFDAKDKLKESED--QLYRYLNYQKHGYILSGKWLRYDKSKVL 174
Db 119 -NDFTILLESKAYNTEIDNK-KVKDNPHFQILRYLGNLKNYKGYFLTNGRFRWFDNS-IL 175
Qy 175 YGKRYIEFNFYKIEBEKEYEQEWFVLYLIRKERYLKTNSVIEVEKQIAKEKILQ 234

Db 176 NSNKVFEINLEKIIEDQNI---EAFAYFVSFSAFNTEKEDHLEITLQNNKLSKIKIE 232
 QY 235 KTLKEILVERPDSIVFK-IAKNIYDKFKLGGKEITQHILASILEESIIFILRIFFIAY 293
 Db 233 DDLKSIIVGTNGNSLFFIGSRIYNK-----TKADLKLIYENSILYFIFRLLFIAY 283
 QY 294 IEDNDIFKKILOENKLYRSSIFRYFF--YDENTKKKLEYKKIIFLNLKDGSDAIKPP 351
 Db 284 FEDK--FEILLEKHYPKSKISLRTLENNLQDESSGGFGELENIYFNKNGKGNFDM 341
 QY 352 VNGGLFSDKVKYLNNEGLLSISEIEILVKMLPFEKNIKDEKFKVYKSLRDPKSGFEL 411
 Db 342 VNGGLFSDKSTALLSTPKIFNDKDLKFTLNQLNFKDKNLSFKR--DYKTLVSVEHLGTI 399
 QY 412 YETLLEYDLRIADTTVHRIIDGVVILIRTEBELE-----NKKVNIATYVL 456
 Db 400 YEGLLSYFPEIAN-----EDIYVSYKEKSKIECYVDNVDYFKILEKSKVKEKTYFYK 452
 QY 457 KGNILYTSRSLDRKSGAYTTPDDLTDFMVISSIEBQKTKSPDKIKIINDSCSGHFLI 516
 Db 453 KQVILKNTSNRKSASAYTPQSTANFLIQSALDKLNNENILAFKILDNACSGHFLV 512
 QY 517 SCLDYLTEKVVYELDKFEDVKELDEEYRVIIIESEYDVQD-SISKELVLKMLLKCI 575
 Db 513 GVLNAITHIVLSDFOHTKLELYBEEKENILNHIKDF-VQDYEVDESILKRLLLKRII 571
 QY 576 YGVDINPISVEITMLSLMINTPIGTPLSFIEHHIKAGNALGYTKDEFFDIVKVKFSG 635
 Db 572 YGVDLNPFSEILTKLSWDSFIPTPLSFIEHHIKCGNALNSNLSDFKDLIK---QNS 628
 QY 636 FSLFKKRIKEIITILEDIYQIKINGDITTKEDIEKSKIKIYKESKIDINLRIIFSILK 695
 Db 629 SNLFTNSITQBEILQEVEPKLDNLKOTNEBQIKSQKIYQN-ELTPKLNKLNLVINYIN 687
 QY 696 LYSLSFDKSLNMFSDIASVILIENLGNKTTSSDEKIEKIRKLSYKFFHYGIBFP 755
 Db 688 --TLHP---VNBELQILKALS--QDDIQNLQNEQAKAI--ISKYKEFNFNFELEFP 738
 QY 756 DIQE-----GPDIVGNPWEKTKNETEFPSKHIPNRYKLGIKQNIKQBIILSKDHP 810
 Db 739 EIVENQVFKGFDIIGNPNMDTKFSDSDFPQYKSDYRSLSASKKEIQDNLAKDY-- 796
 QY 811 LSIYNEEKNSIIANNIYKPFKCTSGDNLFPYFVTNKLKIEKGNITLYVPSAI 870
 Db 797 IKQTEKQKAYINDISEYTKKAYPLNKGSGDGNLFLFVEKLSLLKQDGNLAYVPSAI 856
 QY 871 WNESSRILRKHIFARYKLNIYQFENKKR-FKDVHSSFKFAIFOLSNIKESTSSFKAF 929
 Db 857 MFEDGSLILRKEILENKTLEYFYSFENNKAIFIDVHRSYKFAIMLIK-N-TQANHHTIKM 915
 QY 930 MIQSSD-NILKEITRDLKSDKDAYKGIELNINOIKLSPIQESIIEPKDNEEFTLINKM 988
 Db 916 MFYKTDINSLK-----NKDEI---LTLNLKDKICUSPTTHALMELKQALBILRS 964
 QY 989 FSKFSALGSGYIDFKGLDPSIKNRKSLKCECNKNIILYSGANTHQNSRF-----PE 1043
 Db 965 YNAFQNLSDYIDFRFLD--WINDKDLFIEFREGLLPLEYEGKMLHQFDANFSQATYFL 1022
 QY 1044 DK---DAKESKLLMIDKDEKVLIT-KDNQYOTE--RVFYRAIASNTNERTMISTLSPG 1097
 Db 1023 ERAKFDERLKSLEYRAKATGKELNPKLIKYDRFPFRIGRYKIASDNTNERTLIASLLPK 1082
 QY 1098 NCVCNSIYINDEK-----TPISLYKK-LFIISIFNSFVDFPLRRFVDSNVLKS 1146
 Db 1083 NCGGADSTYNTPKQYVLKDDVICMDIVPEYRILFVLALFNSLWVDFIIRNNVQINVSKS 1142
 QY 1147 CLYQCPMPQPEKEILSNPLNLAKNTSLLIAKNDPENFKYLLVLEYKPKDEKVN--- 1203
 Db 1143 YLERIPLPQSPDEEQNNEIYKTLAKNALLQIYNDON-----RHPDELKQFENIKN 1194
 QY 1204 -KILKLDKDEFFPEKENEENNFIIASLYSLAKEDFITLLGDPKALNKKKGEDYISSL 1260
 Db 1195 BEIPKTKAYDILRAK---NDLLVKELYGLSDSEFSYMTISFKVLNEKQ--SEYITLL 1247

RESULT 7
 QSHV91 CAMJR
 ID Q5HV91 CAMJR PRELIMINARY; PRT; 1250 AA.
 AC Q5HV91;
 DT 10-MAY-2005 (TREMBlrel. 30, Created)
 DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
 DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=CJ0E0789;
 OS Campylobacter jejuni (strain RM1221).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Campylobacteraceae; Campylobacter.
 OX NCBI_TaxID=195099;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX PubMed=15660156; DOI=10.1371/journal.pbio.0030015;
 RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Raeko D.A.,
 RA Ravel J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
 RA Dodson R.J., Durkin A.S., Madupu R., Sullivan S.A., Shetty J.U.,
 RA Ayodeji M.A., Sivarubeyn A., Schatz M.C., Badger J.H., Fraser C.M.,
 RA Nelson K.E.;
 RT "Major structural differences and novel potential virulence mechanisms
 from the genomes of multiple Campylobacter species.";
 RL PLoS Biol. 3:72-85(2005).
 DR EMBL; CP000025; AA034576.1; -; Genomic_DNA.
 DR TIGR; CJ0E0789; -;
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
 DR GO; GO:0006306; P:DNA methylation; IEA.
 DR InterPro; IPR007409; DUF450.
 DR InterPro; IPR002296; N12N6_mtfase.
 DR InterPro; IPR002052; N6_Mtase.
 DR Pfam; PF04313; HSDR_N; 1.
 DR PRINTS; PR00507; N12N6MTFRASE.
 DR PROSITE; PS00092; N6_MTASE; UNKNOWN 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 1250 AA; 147549 MW; A367AD74887B235 CRC64;
 Query Match 27.0%; Score 1764.5; DB 2; Length 1250;
 Best Local Similarity 34.1%; Pred. No. 6.1e-60;
 Matches 452; Conservative 250; Mismatches 474; Indels 149; Gaps 41;
 QY 5 DIVTNNPNISLYQLSKDFIKKENINKLKDPF-----ILIKNKLFSIDDDNSTEAN 55
 Db 6 ELKYLNNLFTSYSLDIDFPKIYNFNANQAFRDLTKITKYDNKNFKQNEHQFE-- 63
 QY 56 IESLLKYIFEELNYSVEQKAGQIEGVESRVDILLFENDKDKASFNNKLKBAKKNNEPIP 115
 Db 64 -DEFISKVLEILGWCFFVRQDEKILQGLEKDPDFLLFSNDKLKSKYENLDKETKSS--- 118
 QY 116 IEDILITAEVKRPTFSFPAKDKLKESED-QLYRYLNQYQKHGYGLNSGKWLRYDKSKVL 174
 Db 119 -NDFTIILSKAYNIENDK-KVQDNPHFQILRYLGNLKNKYGFLTNGRFWRFYDNS-IL 175
 QY 175 YGERKYIEFNPXKIEEKEEYEQEFVLFIYLIRKERVLTNSVIEVEKEQIAKEEIIQ 234
 Db 176 NSNKVFEINLEKIIEDQNI---EAFAYFVSFSAFNTEKEDHLEITLQNNKLSKIKIE 232
 QY 235 KTLKEILVERPDSIVFK-IAKNIYDKFKLGGKEITQHILASILEESIIFILRIFFIAY 293
 Db 233 DDLKSIIVGTNGNSLFFIGSRIYNK-----TKADLKLIYENSILYFIFRLLFIAY 283
 QY 294 IEDNDIFKKILOENKLYRSSIFRYFF--YDENTKKKLEYKKIIFLNLKDGSDAIKPP 351
 Db 284 FEDK--FEILLEKHYPKSKISLRTLENNLQDESSGGFGELENIYFNKNGKGNFDM 341
 QY 352 VNGGLFSDKVKYLNNEGLLSISEIEILVKMLPFEKNIKDEKFKVYKSLRDPKSGFEL 411
 Db 342 VNGGLFSDKSTALLSTPKIFNDKDLKFTLNQLNFKDKNLSFKR--DYKTLVSVEHLGTI 399
 QY 412 YETLLEYDLRIADTTVHRIIDGVVILIRTEBELE-----NKKVNIATYVL 456


```
Db 400 YEGLLSYFPEAN-----EDIIYVYKESKEIECYFDNDYDKILEKSKKVEKYFYK 452
Qy 457 KGNITLRSRDLRKKSGAYTTPDDLTDFWVISSIEEQKTKSPDLIDKIDNSCGSHFLI 516
Db 453 KGQIYLNKTSNSRKASASFYTPQSTANFLIQSALKDNLNENILKFKILDNACSGHFLV 512
Qy 517 SCLDYLTBKVWVYELDKPDVKELDEBVRVIERSEEDVDOD-SISKELVLKMLKKCI 575
Db 513 GVLNAITHVLSDFHFTNLKELYEKEENILNHKOP-VODEYVEDSDIILKRLIKRII 571
Qy 576 YGVDINPISVITMLSLMNTFIFCTPLSFTEHHIKAGNALLGYTKDFFDVIKKPBG 635
Db 572 YGVDLNPISIELTKLSLWIDSFIPTPLSFTEHHIKAGNALLGYTKDFFDVIKKPBG 628
Qy 636 FSLFKRIKEITILEDIYQIKGINTTKBEDIEKSKKIYKEYESKDIDLNLRIIFSLIK 695
Db 629 SNLFTNSITQFEILQVFEKLDNLKOTNEBQIKSQIYQN-EITPKLKLNLVLYNIN 687
Qy 696 LYSLSFKSLNMFSDIASVISL--IENILGNKTSSEDEKIEKIRKLUSYKFFHYGIE 753
Db 688 --TLHF---VNKEELQILKALSQDDIQLN-----SQNEQAKAVISKYQKEBFFNFYELE 736
Qy 754 FPDIOE-----GPDIVIGNPWEKTKENETFEFSKHIPNRYKLGITKEQNIILKQILSKDN 808
Db 737 FPEIVENQVFGFDIIIGNPPWDKTKFSDDDFFPOYKSDYRSLSIAASKKEIQDNLAKDY 796
Qy 809 HPLSTEYNEEKNSITAINNIYKDFKCTSGDNLFPYFTFNLKIKERKGNITLYLPS 868
Db 797 --IKQNYEKQAYINDLSEYKAYPLNKGSDGNLFLVEKNLSLLQDGNLAYVLPS 854
Qy 869 AIWNESSRIIRKHIFARYKLNIIYQFENKCR-FKDVHSSPKFALPOLSNIKESTSSPKA 927
Db 855 ALMPEDGSLTLRKEILENKTLLEYFYSFENRQAFADVHKSFKALMQVKN-TQANHTHKI 913
Qy 928 KFMIOSSD-NILKEITRDLKSDKDAYKGIELNINQIKLSPIQESIIEPKDNEEFTLIN 986
Db 914 KTMFYKTDNPLK-----NKDEI---LTLSLKDIKIKLSPTLALMELKDKQALEILR 962
Qy 987 KMFSPFSALGEGYIDFKKGLDPSIKNRKSLKCECNKNKLIPLYSGANTHFNRSR- 1041
Db 963 KYNFAQNLSPYIDFVNL--HWINDKDLFIEEFREGLLPLFGKMLHQDTNFSQATY 1020
Qy 1042 -----FEDK-DAKESKLLWIDKEDLEKVLTKDNOYQTE--RVFYRAIASNTERTMIS 1092
Db 1021 FLEKAKFDERLKSLSRAKATGKELNPKLIK---YDREPERLGYRTIASDTNERTLIA 1077
Qy 1093 TILSPGNCVCVNSIYINDEK-----TPISLYKK-LFISIFNSFVDFELLRFRVDS 1141
Db 1078 SLLPKNCGCGNSVYGNIPKQYIVKDDAICMDIVPYERILFVLALFNSLWVDFIIRNMVQI 1137
Qy 1142 NVLKSCLYQCPMPQPEEKEILSNPLYLNLAKNTSLLIAKNDPENPKYLLYLEYKPKDKEK 1201
Db 1138 NVSKTYLERIPQPSDEIQNEIYKTLAKNALLQLVNDKQH-----HFDLQKE 1189
Qy 1202 VN---KILKDKDEEFFKEKENENNFIASISLAKEDFITLLGDFKALKNKKGGEDYI 1257
Db 1190 FNIKNEEIPKTKAYDILRAK---NDLLVKKLYDLSDDFEFSWISTFKVL-NEKQSE--Y 1243
Qy 1258 SSLIK 1262
Db 1244 TTLK 1248

RESULT 8
O50806 BORBU
ID O50806 BORBU PRELIMINARY; PRT; 422 AA.
AC O50806;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein BBK02.
GN OrderedLocusNames=BBK02;
```

```
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Plasmid lp36.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=139;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=96065943; PubMed=9403685; DOI=10.1038/37551;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
RA Winn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A.,
RA Artiach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi";
RL Nature 390:580-586 (1997).
DR EMBL; AE000788; AAC66148.1; --; Genomic_DNA.
DR PIR; D70251; D70251.
DR TIGR; BBK02; --.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 422 AA; 49766 MW; 2A87FB56008EFEAB CRC64;

Query Match 25.2%; Score 1648.5; DB 2; Length 422;
Best Local Similarity 75.7%; Pred. No. 5.4e-56;
Matches 315; Conservative 49; Mismatches 49; Indels 3; Gaps 2;

Qy 853 LKLIKEKGNLTLYLPSAIWNESSRIIRKHIFARYKLNIIYQFENKCRFKDVHSSPKFAI 912
Db 1 MKLVKPGGNLTLYLTPSSLWSESGSKALRQHIFPSNKLNIYIYQFNQKGFRAWTGFKFAI 60
Qy 913 FOLSNIKESTSSPKAFKFWMIQSSDNLKBITRDLKSDKDAYKGIELNINQIKLSPIQES 972
Db 61 FOINSKEPTKFRVKFVIQSSDNIKEITSDKAGENAYKGIELDIAQIKLSPIQES 120
Qy 973 IIEPKDNEEFTLINMFSPKFSALGEGYIDFKKGLDPSIKNRKSLKCECNKNLIFLYSGA 1032
Db 121 IIEPRDSAEFTLVNKMFSRPTLIQYIDFREGL--NLTKYKALYKEYNEKFIPLYSGA 178
Qy 1033 NIHFNSRFPEDKDAKESKLLWIDKEDLEKVLTKDNOYQTEVFYRAIASNTERTMIS 1092
Db 179 NIHFNSRFPEDRAAKESKLLWIDKEDLEKVLTKDNOYQTEVFYRAIASNTERTMIS 238
Qy 1093 TILSPGNCVCVNSIYINDEKTPISLYKKLIIFISIFNSFVDFELLRFRVDSNVLKSCLYQCP 1152
Db 239 TILSPANCVCVNSIYINDEEIPISLYKKLIIFISIFNSFVDFELLRFRVDSNVLKSCLYQCP 298
Qy 1153 MPQPEEKEILSNPLYLNLAKNTSLLIAKNDPENPKYLLYLEYKPKDKEKVKNLKDKED 1212
Db 299 MPQPEEELNLSLYLTAKNTSLLIIVKNDPDNFKYLLYLEYKPKDKEKVKNLKDKED 358
Qy 1213 EFPKEKENENNFIASISLAKEDFITLLGDFKALKNKKGGEDYISSLIKGYDNYL 1268
Db 359 EFPKEKENENNFIASISLAKEDFITLLGDFKALKNKKGGEDYISSLIKGYENYL 413

RESULT 9
Q5XYI6 BORBU
ID Q5XYI6 BORBU PRELIMINARY; PRT; 149 AA.
AC Q5XYI6;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=EGP295;
OS Borrelia garinii PBi.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=290434;
RN [1]
```

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBI;
RA Glockner G., Lehmann R., Romualdi A., Pradella S.,
RA Schulte-Spechtel U., Schillabel M., Wilske B., Subnel J., Platzer M.;
RT "Comparative analysis of the Borrelia garinii genome."
RL Nucleic Acids Res. 32:6038-6046(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBI;
RA Gloeckner G., Schillabel M., Lehmann R., Platzer M.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY729404; AAU86146.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 149 AA; 17344 MW; CB660C79D86F038E CRC64;
Query Match 10.3%; Score 671; DB 2; Length 149;
Best Local Similarity 85.9%; Pred. No. 7.3e-19;
Matches 128; Conservative 14; Mismatches 7; Indels 0; Gaps 0;
QY 340 LLDGSDAIKFPVFNGLFSEDKVYLNNEGLSISEIEILVKMLFFFEKNIKDEKFKV 399
Db 1 MLDGSDAIKFGIFNGGLFAEDKVKYLNKSLLSISEELVKLIFFFEKNIKDEKFKV 60
QY 400 YRLDPKSPGELYETLLEVDLRIADTVHRIEDGVYLIRTEELEENKVKNIATYLGKN 459
Db 61 YSKLDPKSPGELYETLLEVDLRIADTVHRIIVKDGVLIRTEELEENKVKNIATYLGKN 120
QY 460 IYLTSLRSLDRKSGAYTTPDDLTDFMVIS 488
Db 121 IYLTSLRSLDRKSGAYTSDLLIYFMVIS 149
RESULT 10
Q4HA54_9DEIO
ID Q4HA54_9DEIO PRELIMINARY; PRT; 1318 AA.
AC Q4HA54;
DT 13-SEP-2005 (TREMELrel. 31, Created)
DT 13-SEP-2005 (TREMELrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMELrel. 31, Last annotation update)
DE Similar to Type II restriction enzyme methylase subunits.
GN ORFNames=DgeODRAFT_1104;
OS Deinococcus geothermalis DSM 11300.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=319795;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Istrani S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHE01000004; BAL83002.1; -; Genomic_DNA.
KW Methyltransferase.
SQ SEQUENCE 1318 AA; 147484 MW; 2C5E0B8D91D73358 CRC64;
Query Match 9.3%; Score 605.5; DB 2; Length 1318;
Best Local Similarity 20.2%; Pred. No. 2.7e-15;
Matches 259; Conservative 247; Mismatches 471; Indels 305; Gaps 52;

QY 148 YLNOYQKH-YGILNGKVRILY-DKSKVLYGEKRYIEFNFXKIESEKYEKQEWFLFIY 205
Db 145 FLNRSGEHTWGLVNGERLRLRDAQVT--RPAFVEFD---LEBMFQDGSAGFVLWL 199
QY 206 LIRKERYLKTNS-VIEVEKEQIAKEKIIQKTELKELYERDDDSIVFKIANNIYDEKPKL 264
Db 200 LTHRSRFEKGGAIIEGWNKQASAGGTANDALRDGV-----ERAIAITLGSGLFNRDL 254
QY 265 SGKEITQHILASILEE-SIIFILRFFIAYIEDNDIPKKILQENKLYRSSISFRFFPYDE 323
Db 255 CAQLATGERQADLYRLSLKLVYQLIFLVVEDRDLLHPADTD-----PTAKARVAHST 309
QY 324 NTKKL-----EYKKIITIFNLLDKGSDAIKFPVFNGLFSEDKVYLNNEGL 371
Db 310 RRLQIAEASGSATHGDWEGLVKLLRGMVGLPFLGPAFGHLPFQYPLLAGC----- 365
QY 372 LSIISIEILVKMLFFFEKNIKDEKFKVYKSLDKPSGELYETLLEVDLRIADTVHRII 431
Db 366 LANSDLVAAIRALSIEDGGTL--KSNFTDLDAAELGSIYESLLELHPEINTAT----- 418
QY 432 EDGVYLIRTEELEENKVKNIATYLGKNIYLTSLRSLDRKSGAYTTPDDLTDFMVIS-- 489
Db 419 -GFTLSS-----AAGN-----ERKTTGSYITPTGLLELLESSUD 453
QY 490 --IEEQKTKSPL--DIKIIDNSCGSGHFLISCLDYLTBKVWYELDKFE-DVKKELDE 542
Db 454 PVIEDALTKPDVAALKALNVVDPAACSGHFLAA---ARRIGLALARAHDVTPQSP 509
QY 543 EYRVIIIESEEDYVQDSISKELVLRMLLKIKCIYGVNDINPISVEITMLSWINTIFGTP 602
Db 510 QLRAATRE-----VIAHCYIGVDLNPMAIELAKVALWLESAPVGP 550
QY 603 LSFTEHHIKAGNALLGYTKDBFFDIVKKKPSGSLFKPKRIKIITIL-----EDIYKI 657
Db 551 LAFIDHRURSGNSLGTTP-----DLMRREDSIPEKLVRRVLPKPAVKAVTLHLDPAPAE 606
QY 658 KGINDTTKEDIKSKIYKE-----YEESKDIDLN-RIIFSL--IKYLSLSPD 702
Db 607 EGDDKTVRALKNRNKQEREDLIERAGSGQLFDLQRLDLSGLQRMVEALDRIEPLDSLHV 666
QY 703 KSLNWEFSDI-----ASVISLIENILGNKTSSEBKE 733
Db 667 QAQESTWSAQIEAHHRRLKTLADAWCAAFVTPKTPGAPAITTRTLQDLDRNPDSLNT 726
QY 734 KIEKIRKLSYVKFHYGIEFPDI--OEGFDIVIGNPPWEKTKFNETEFFSKHINVRK- 790
Db 727 AL--VGLAEQYRFFHWHVEFPDVAQGGFDCVGNPWEKIKQEKWEFAARVEPIAAA 784
QY 791 -----LGIK-EQNIKQEIILSKDNHPLSIEYNEEKNSIIAINNIYKDFKCTSG 839
Db 785 PNAARTLIALKTEQPAIYADFVR-----DLRQAEGESHFIRASGRYP-----LTGR 833
QY 840 GDPNLFRYVTFNKLKIEKGNLYLVPSATWNSSSRILKHIIPARYKLVNIYFENKK 899
Db 834 GDVNTYAI FSELQCSLNPQGRMGII VPTGIATD TTKFFQSLMDKASLASLYDENRE 893
QY 900 R-FKDVSSEKFAIFQLSNIKESTSPKAKPMIOSSDNILKEITRLDXSDKDAVKGIEL 958
Db 894 GIFPGVHRSFKCLLTVSGEERPNAAQAFAPALNAD-----EARDPG-KVFTL 941
QY 959 NINQIKLSPIQESIIEFKDNEEFTLINKMFSKF-----SALGEG--YIDFKKGLDPSIK 1011
Db 942 SPAEIALNPNTRTAPVFRSSRDAITKGIQYRVPLVNEATGENPWSGFWMFDMDS-- 999
QY 1012 NRKSLKKEC-----NNKNLI FLYSGANIHOFSNRFF-----BDKAK 1048
Db 1000 NDSGLFRTAEQLDADGWELGNRYKGDGQMLPYEAKLMEHQFHRFATYTESGTRDMT 1059
QY 1049 ESSKL-----LWIDKEDLE-KVLTKDNQ-----YQTER---VFYRAIASNTNERTMI 1091
Db 1060 SAEKARADALPLPRYVWPQTEVERDLIQKNGNVIWETWDLWLMGWRDIARTDERTFI 1119
QY 1092 STLSP-----GNCY-----CVNSIYINDEKTPISLYKKLFIISIFNSFVDFLLRRFVDS 1141


```
Db 1120 GAVYPRRAAGDTFLOMLPDC-----PVAQIPGL--LGTLSNFAHDFAVRQKVG 1166
Qy 1142 NVLK-SCLYQCMPQPEEKEILSNPLYNLAKNTSLIAKNDPENPKYL----LYLEYFK 1196
Db 1167 THLYKYNVTRQBPVLP-----SAPTQFORLAFTIPRVLELTYYTA 1204
Qy 1197 PDKVKVKNILKDKEDDEFKEKEN-----ENNFIASLYSLAKEDFILLGDFKALKNKK 1251
Db 1205 HDLAGFADLDGDPFPFVNDNERFWLRAELDALYFHLYGIPRDVDYIMETFPFVRRKD 1264
Qy 1252 KGEDYISSLIKGYDNYLLNNKI 1273
Db 1265 EAQ-----YGYRTKNAI 1277

RESULT 11
ID Q7NGZ7_GLOVI PRELIMINARY; PRT; 1442 AA.
AC Q7NGZ7;
DT 01-MAR-2004 (TREMELrel. 26, Created)
DT 01-MAR-2004 (TREMELrel. 26, Last sequence update)
DE G112740 protein.
GN OrderedLocusNames=g112740;
OS Gloeobacter violaceus.
OC Bacteria; Cyanobacteria; Gloeobacteria; Gloeobacteriales; Gloeobacter.
OX NCBI_TaxId=33072;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=PCC 7421;
RX MEDLINE=22977040; PubMed=14621292;
RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
RA Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimo S.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
RT cyanobacterium that lacks thylakoids.";
RL DNA Res. 10:137-145(2003).
DR EMBL; BA000045; BAC90681.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0008306; P:DNA methylation; IEA.
DR InterPro; IPR002296; N12N6_mtfase.
DR InterPro; IPR002052; N6_Mtase.
DR PRINTS; PR00507; N12N6MTFRASE.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
KW Complete proteome.
SQ
SEQUENCE 1442 AA; 163040 MW; C9FE608F166E85C1 CRC64;

Query Match
Best Local Similarity 22.7%; Pred. No. 3.4e-15;
Matches 332; Conservative 245; Mismatches 562; Indels 324; Gaps 69;
Qy 19 QLSKDFIKKENTK-----LKDFILIKNKLFSIDNTEANIES-LLKYIFEELNY 69
Db 24 ELEKLNREKLKSKATGTGAQALREFDWCYRRKRELGVKSGAIRNQVIDPILPLRGY 83
Qy 70 S-VEQKAGQI-EGVESRDVILLFENDKDKASFNNKLKAQKNPEPIEDILLIAEYKR 127
Db 84 DRLEAAEPVQTRGLEAGHLAI-----SADGSKLRITWTD-----LDIDLDPAPKR 131
Qy 128 -PTFSFDADKLKESDQLYRLNVOQKHGYLSLNGKWRLY-----DKSKVLYGEKRY 180
Db 132 GAAYRF---SHURIAQ-----RVLLACSERVGLITNGVELRLLSLDPAPDSQVIIP----- 180
Qy 181 IBFNFXKIBEKYEKEQWFLVFLYLIRKERYLKTSNVIEBKQIAKEKEIIQKTLKEI 240
Db 181 IDANKRSREVPDSLR-----LLALASPAGKAVPGVLDKARLQQAQVTKDLRQAREA 235
Qy 241 LYRPDDSDSYFKIAKNYDKFKLSKETQHILASILEESIFILRFFIAYIEDNDIF 300
Db 236 V-----ELFQVRLDHPANRE-KLSEFADRARLAKELWREGLIITTYRLLFILKJESSDDP 289
```

```
Qy 301 KKILQ--ENKLYRSSIS-----FRY--FFYDENTKK-KLEYKIIITIFNLIDKGSDAIKF 350
Db 290 ARSFSFASTSLWRNSFSVALARYVRMALDEGAETGQLLERGLRSLFRMFAEGLQCTEL 349
Qy 351 PV--FNGLGFSEDKVKVILNNEGLLSIS--TEEILVKMLFFEEKNIKDEKFKVYKRLDPK 406
Db 350 NIRPLGSLFGAHATPILSD---LAWGERAVAYLLDRLLWTLLKGTVARVHYGPLDVE 406
Qy 407 SFGELYELLBYDLRIADTTVHRI-----IEDG-VYLIR-----TEEELNKK-- 447
Db 407 DLGRVYEALELEPGIADEPMCLRRQKLEVVVPIDQGEKYRVTAAVKNDSEDEAEEDDA 466
Qy 448 -----KVNKIATYLGNIYLTSLDRKSKGAYYTPDLDITDFWVISSIEE 492
Db 467 EESQFEETSGRGKTKVWEIEAIPGRPYLRV-GLGRKASGSYYTPHSFVFLVQETLGP 525
Qy 493 QLKTQSP-----LDIKIIDNSCGSGHELISCLDYLTEKVMYELDKFED----- 535
Db 526 QVARSPOSDFKPLEILKLVCDPAMSGHFLVACRFLGKL-YEARLCEDELAMAAEK 584
Qy 536 -VKKELDEEYRVIIIESEY-----DVQDSISK-----EVLV 566
Db 585 RAESAEDQQRRAALEGAKLFQORVADLPDPNDELLKVLPSRAPGLGLSQRKAEAMC 644
Qy 567 KRMLLKXCIYGVNDINPISVEITMLSLMINTFIFGTPLSFIHHIKAGNALLG-YTKD--- 622
Db 645 RRLIAVHCLYGVQKNPLAVELAKLSLWIESHAEGPLTFLDHLRLVLGSLAGPFLKDLR 704
Qy 623 -----EPPDIVKKKFSFSLFKRIKEIITILEDIYQIKGINDDTKEDIEKSK--- 672
Db 705 YPGSQKQMDLLISGLKEGFT---KALSNALRHVHDLSESV---GATLSEIAKQAAKT 757
Qy 673 KIYKEYE-----ESKDIDNLRIFSLIKLYLSLSPDKSLME---FSD 711
Db 758 RLDKALEPFMIVAAAGGMLSGPECDSDAVARLAHAVASTSALPAD--INKEERLIRM 815
Qy 712 IASVISLIENLTGNKTSSEDEKEIEKINLSYKFFHYGIE--PP-----DIQEGFDIVI 765
Db 816 IAKGMAMQENL---TTIEEVFOQKHIEKPLALPFDLTFFVEVFPFHNSCDYRNGFDVLL 871
Qy 766 GNPWEKTKFNETEFFSKHINRYKLGIGKEQNIKQELTSLKDNH----- 809
Db 872 GNPPWSKSNL-----ELPEY--MGAVDPREL--EVTSSDERIDLESYKSSAWNAM 919
Qy 810 -PLSIEYNEEKNSIIANNIYKDFPKCFTSGDGNLFRYFVTNFKLKEKGNLYLVP 868
Db 920 LLASLEEDRMCKIISLLNPDAQSDLVGFGS-GHADIFASFDVRLFSLLRVNGLFAMVLP 978
Qy 869 AIWNSSRILRKHIFARYKLYNYIQFENKRRKFDVHSSPKF--AIFOLSNIKESTSSPK 926
Db 979 SLHVSQSLRRLRRKLLNKAELISYFSPENKYSLEFHKSWKFTPVIIIRNSEPKER-KRPP 1037
Qy 927 AKFMIQSSDNLKKEITRDLKSDKDAYKIELNINQIKLSPIQESIIEFKDNEFTLIN 986
Db 1038 AQFYLHDSWLFS-----NKDPYYLYDPNIEVMD-----SESLI-----FQBYTSVS 1082
Qy 987 KMFKFSALGSGYIDFKKGLDP---SIKN-----RKSLLKEC-----NNKNLIFY 1029
Db 1083 DMKCAETIFGNS-LSWKYKSENTYWSIRRELNATDWRVIEYSPKCGRTWSKDEALIH 1141
Qy 1030 SGANTHQFNRPFFEDKAKESKLLWIDKEDLEKVLTKDNQYQTERVF-----YRA 1080
Db 1142 QPGTIHQYTD-----LWTGAKTL--TIPFNLYDRPVLELSRYFAAAYRM 1185
Qy 1081 IASNTNETMTSL-LSPGNCYCVNSIYN--DEKTPISLYKKLFIISIPNSFVDFILRR 1137
Db 1186 SARATDERTSITILTPTGTT-ATNSLPTEGCPQKRPNRL--SLGAVAVCNSSVFDWYLR 1242
Qy 1138 FVDSNVLAKSLYQCPMPQPEEKETILSNPLYNLAKNTSLIAKNDP--ENPKYLLYLEVF 1195
Db 1243 RVGSKTISKFI-----MONTPISDLASQN---NLCVHSSRLVSNHFGYETLWHEQVCDEW 1295
```

```
QY 1196 KPDKEKVKILKDKEDSFFKEKENENNFIASLYSLAKEDFITLLGDFK----- 1245
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1296 REPKEPFTVPV-LESDEEM-EVRSADAVVADAGLEREQYAHILSFSHRSPYKPKL 1353
QY 1246 -ALKNKKGGEDVISSLIKGYDNY 1267
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1354 CLAKPELQELGIGLGAFTKKYDFY 1376

RESULT 12
OS2000 HALSA
ID O52000 HALSA PRELIMINARY; PRT; 706 AA.
AC O52000;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=VNG5138C;
OS Halobacterium salinarium (Halobacterium halobium).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1];
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NRC-1 / ATCC 700922 / JCM 11081;
RX MEDLINE=99063795; PubMed=9847077;
RA Ng W.V., Clufo S.A., Smith T.M., Bumgarner R.E., Baskin D., Faust J.,
RA Hall B., Lorez C., Seto J., Slagel J., Hood L., Dassarma S.;
RT "Snapshot of a large dynamic replicon in a halophilic archaeon:
RT megaplasmid or minichromosome?";
RL Genome Res. 8:1131-1141(1998).
RN [2];
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=NRC-1 / ATCC 700922 / JCM 11081;
RX MEDLINE=20504483; PubMed=11016950; DOI=10.1073/pnas.190337797;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA IseNBarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AF016485; AAC82880.1; -; Genomic_DNA.
DR PIR; T08313; T08313.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR InterPro; IPR002296; N12N6_mtfrase.
DR PRINTS; PR00507; N12N6MTFRASE.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 706 AA; 80793 MW; 16D0E0F72A867B2 CRC64;

Query Match 8.5%; Score 558.5; DB 2; Length 706;
Best Local Similarity 24.1%; Pred. No. 8.7e-14;
Matches 180; Conservative 150; Mismatches 219; Indels 199; Gaps 32;

QY 62 YIFEELNYSVEQKAGQIGVESRDVIL-LFENDKDKASFNNKLKEAKKNRP-----I 114
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18 YIDEHLPEFEEWDEVSDELQAYDDIMALWREKDTAPKNESQLEEFIRPMFKLGI 77
QY 115 PIEDILIIAE-----VKRTFFSP-----DAKDKLKESED----- 143
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
78 PFE-----VEESTSRQRRDYGFEDDAARAFERREGGDFYENAVAVADAKRWGRL 133
QY 144 -----QLYRYLNQYQKHGYILNSGKVRWLVDKSKVLYGE-----KRYI 181
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
134 DTRGSGEHRDPENPSYQIHVLTQETPARWAVLTDGKKWRLY-----YGPTRHLDYSY 187
QY 182 EFNFXKIEEKEQEWFVLFTYLIRKRYLTSNVIEVEKEQIAKEIKLTKL 241
```

```
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
188 EVDLPITLESGLDEDFKVFYLF-----FRHEAFLDGSG-DCFLDDVDYDSNVFAQELGDL 243
QY 242 YERDDIVFKIKNIVDKFKLSKEITQHIILASILEESIIFILRIFFIAYIEDNDFK 301
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
244 QDN-----IYBAIKILSEGVLQYPENDLDEGDLIHDSSLIYLRFLVLYAEAG--R 296
QY 302 KILQ-ENKLYRSSISF-----RVFFYDENTKKKLEVKKIITIFNLDKG 344
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
297 ELLDTSNEIYEQSYSLNSLKQEIABELDSGPKTRDQDNLOSLD-----ELFTLNKG 351
QY 345 S-----DAIKFPVFNGLFSED-----KVYLANNE--GLLSISIEBILVKMLPFE 388
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
352 SKSRGIPEDLYIPAYNGGLFRTDPDEDDSAEAKFLANHDVDGDAYLAKVWELLTR---S 407
QY 389 EKNIKDEK-FVKYSRLDPKSGELGYETLLEVDLRIADTTVRIIEDGYLIRTBEELENK 447
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
408 KNDGGGKIFVDYSSLDVRHLGSIYEGLLYQLNVADEPL--ALDDGGEV--SADEGDV 463
QY 448 KVNKIATYLVKGNLYLTSRLDRKKSGAYTTPDDLTDFMV-----ISSIEEQLTKSP- 499
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
464 VVQ-----EGEYLTDTDSGERKATGSIYTYEYVYIVENTLPLVDVDIRADLAGRSAR 517
QY 500 -----LDIKIDNSCGSHFLISCLDYLTETKRYWYELDKFEDVKKELDEBYR 545
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
518 GEDRGPAEFAERIFDLKILDPAMSGHFLTSAIDYLAREI-----IDAQER 564
QY 546 VIIIE-----SEEDVQDSISKEIVLKMMLKXCIYGVNDINPISVEITMLSLWNTPIFG 600
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
565 QAAQGIETVQNDHDIINWA-----RRQVQRCTCYGVDLNPLAVELAKVLSLWLTAAE 617
QY 601 TPLSFIEHHIKAGNALLGYTKDEFPDIVKKK-----PESGSLFKKRIKEIIT 648
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
618 QPLAFDLHHLTKGSLVGSNVVE--IDALERNGAAAEAGPNSTLIDFGMT-FEGTINELWA 675
QY 649 ILEDIYQIKIGINDTTEDIEKSKKIYK 676
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
676 IHADLI-----AIENELEEDVKAMEKKYQ 699

RESULT 13
QSHI06 HALSA
ID QSHI06 HALSA PRELIMINARY; PRT; 662 AA.
AC QSHI06;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Vng6135C.
GN OrderedLocusNames=VNG6135C;
OS Halobacterium salinarium (Halobacterium halobium).
OC Plasmid pNRC200.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1];
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=NRC-1 / ATCC 700922 / JCM 11081;
RX MEDLINE=20504483; PubMed=11016950; DOI=10.1073/pnas.190337797;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA IseNBarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005149; AAG20810.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR InterPro; IPR002296; N12N6_mtfrase.
```



```
QY 758 QF---GFDIVIGNPWEKTKFNEBTFFSKHINRYRKLGIKEQNIILKQIBLSK--DNHPLS 812
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
746 MERDGGFDVIGNPWERIKLQEQFFAARSPEIATPNKAE--RQKLNLEADPAS 802
QY 813 IYNEEKSGIIANNIYKDFCKFTSG-----GDPNLFYVTFNKLKIKEKNLTVL 865
Db :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
803 ADGRWLQDFIFAKRTAEASSEPARTSGRYPLAGRGDVNTYALFASLFRSLTKPEGRAGVI 862
QY 866 VPSATWNESSRILRHIFARYKLVNIYOPENK-KREKDVHSSFKFAIFQLSNIKES--- 921
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
863 VPTGIATDYTLRFLFSDLLKRLVQALMFDEWKLFPSPVDPDTFGFVFGFTKEAPIF 922
QY 922 -----TSSFKAKFMIOSSDNILKBITRDLKSDKDAYKGIELNINQIKLSPIQES 972
Db :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
923 CAVALCIEHALDPRRFLMSPA-----ELAR-----INPNTKTAPVRS 961
QY 973 -----IIEFKDNEFTLKNMFKSPSALGEGYIDFKKGLDPSIKRKSIL 1016
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
962 RADAELTAKLYNHAPVLIEKPEEDGGDINPWGITFQTMFHMNSDSPFLTPAQMEAGW 1021
QY 1017 LK-----ECNNKILIFLYSGANIHFNSRFP-----EDK-----DAK----- 1048
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1022 RREGTDVWRETAGSERRVPLYEAKMIHQFDRWASYGNGDTEEDKAQDCTLDEKONPDF 1081
QY 1049 ESKLLWIDKED-----LEKVLN-----KDNQYQTERV 1076
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1061 -----LEKVLN-----KDNQYQTERV 1076
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1142 WLAGAIPALERHAPQEAIDITLVGRNQDWRGAPKVSLEFLIDPKTVATQBMQRETPLV 1201
QY 1077 F-----YRAISNTNERTWISTLSP-----GNCVC 1101
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1202 LSDMLLIAEGSDDAELGELLVTAQOPRLMGWRDITNATNERTWIASVFPKVGXKIP 1261
QY 1102 VNSIYINDEKTPISLYKFLPISFNSFVDFLLRRFVDSNLKSLCYOCPMPQPEKEI 1161
Db :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1262 LMNL-----SEIEPKAAALACLSLTCDVAKQKIGGTTINFLYR-----QLAI 1308
QY 1162 LSNPLYNLAKNTSLIAKNDPENFKY-----LLYLEYKFKDKKVNKILKDKED 1212
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1309 LS-----PSEFSFSLHFIIVPRVLELY-----TSHVMKLAED 1342
QY 1213 -----EPFKEKENENNFIASL-----YSLAKEDFITLLGDFALKNNKKGEDISSL 1260
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1343 LDYSGPPFAWNSRRATLRAELDAFFARQYGLTRDELRYLDPVDV-----KGTDPSET 1397
QY 1261 IK 1262
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1398 FR 1399

RESULT 15
Y3402 METJA
ID Y3402 METJA STANDARD; PRT; 1181 AA.
AC Q60301;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical adenine-specific methylase MJBCS02 (EC 2.1.1.72).
GN OrderedLocusNames=MJBCS02;
OS Methanococcus jannaschii.
OG Plasmid small ECE.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcales; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
```

```
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii."
RL Science 273:1058-1073(1996).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA adenine = S-
CC adenosyl-L-homocysteine + DNA 6-methylaminopurine.
CC -!- SIMILARITY: Belongs to the N4/N6-methyltransferase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; L77119; AAC37060.1; -; Genomic_DNA.
DR PIR; B64516; B64516.
DR TIGR; MJBCS02; -.
DR InterPro; IPR011639; Eco57I.
DR InterPro; IPR002296; N12N6_mtfase.
DR InterPro; IPR002052; N6_Mtase.
DR Pfam; PF07669; Eco57I; 1.
DR PRINTS; PR00507; N12N6MTFRASE.
DR PROSITE; PS00092; N6_MTASE; 1.
DR Complete proteome; Hypothetical protein; Methyltransferase; Plasmid;
KW Transferase.
SQ SEQUENCE 1181 AA; 138439 MW; 9F961D8F8C6A4BDD CRC64;

Query Match 8.1%; Score 526.5; DB 1; Length 1181;
Best Local Similarity 22.3%; Pred. No. 2.6e-12;
Matches 322; Conservative 212; Mismatches 399; Indels 509; Gaps 76;

QY 25 IKKENINKLQDFILIKNKLFSIDD-NSTEANIEBSLLK---YIFELNYSVEQKAGQI- 79
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15 IPKEVYNKKDFIKLVGNIIISIQIPELEKFKPELLKGRYFNKSDDF-----GGQLP 68
QY 80 -----EGVESRVDILAFENDKDKASFNKL-KKAKQNEPIP-----TEDILIIAEV 125
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
69 EPFTROKVIPIELFELGYE-----FTSEISKSPGLDKRIPDYRVSVFNKEILIEAE- 120
QY 126 KRPFTS-PDAKD-----KLKESDQLYRYLNOYQKHGYILSKGWRLYDKSKVLGEXRY 180
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 -----PLGSDLNKXDSGHQVKE-----WLIKSYGVDVTGATNGLEWVLLHYDDTIKEIRTL 173
QY 181 IEFNPKIEE-----KEEYKEQEW---FVLFIYLIRKERYLKTNSVIEVEKEQLA-KEK 230
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
174 KELNLKSIFFVLENKKDKDLLENELKQVSEFYCFSE-YIE--EYIEVATKNIKHKKE 230
QY 231 EIIQKTLKELIYERPDSDIV--FKIAKNIYDKFELKSGKEI-TQHILASILEE----- 280
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231 EITNEFYKEFV-----KLVPFGEDVKDKKSSSEKDKGTGKCLYNCIEAPNTSEL 284
QY 281 -----SIIFILIRPIAYIED-----NDIFKKLQENKLYRSSISFRYF-----FYD- 322
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
285 DKKFAVALLMRLIFIKFLEDKIGIVPRDLRLRTYEDYK--KSNVLINITYDAYLKLPLFVEV 342
QY 323 -----ENTKKKLEKKIITIFNLDDKGSDAIKFPVFNGLGFSEDKVKVYLNNEGLIS 373
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
343 LNPEDERKENIRTPYKDI-----PVLNGGLFRSNNVP---NELSFT 383
QY 374 ISEIEELVKMLFFEEK--NIKDEKFKVYGRLPDPSKSGELYETTLLEYDLRIADTTVHRI 430
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
384 IKD-NEIIGEVINFLERYKFTLSTSEGSEVELNDILGYVE----- 425
QY 431 IEDGVLLIRTEELEENKKNKIATYLNKNIYLTSLDRKSKSGAYYTPDDLTDFWVSSI 490
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
426 -----KLINILAE--KG-----OKGLGAYYTPDEITSYAKNTI 457
QY 491 E-----EQLK-----TKSPLD-----IKIIDNSCG 510
```

Db 458 BPVVVERFKBIKNWINDINFSTDEILNEDSKIAENKHILRAFLDELDIRILDPVAVG 517
QY 511 SGHFLISCLDYLTKEVWVYELDFEDVKKELDEYRVIIIEESEEYVDQDSISKELVLEKML 570
Db 518 SGHFLISALKELLQ-----IKKRI---YLLREENDIY-----KE---KLGI 553
QY 571 LKXCIYGVNDINPISVEITWLSWI-----NTFIFGTP-----LSFIEHHIKAGNALLGYTK 621
Db 554 ILNLYGVNDIDDIJAEIAKLRLALAIENLDVEALKRGEVLLPNIEYVNRVCGNSIAGWID 613
QY 622 DEFFDVKKKPESGSLFKRIKEIITILEDIYQIKGINDTTKEDIKSKKIYKEYEES 681
Db 614 ENL-----KQISISYLCNVR---IMCVLEGLI-----INAHNSEERKKLKA-KELLEK 659
QY 682 KD---IDNLRIFIISI-KLYSISFDKSLNM-----EFSDIASVISLJENIL----- 723
Db 660 RDGYVLDNVEAYHLLYEVRTSHGLKANLLKELDEIRD-----SIYESVTPPAYPAIY 714
QY 724 --GN-----KTSSEKKEIKIRKLSYYKFFHYGIBFPDI--OEGFDIVIGNPPWEKT 773
Db 715 QNGNNKNGKSKKRNPRVEFEKL-----KPFHWKIDFGWIIKEEGFDVIGNPPYGNL 770
QY 774 KENETEFPSKHPNRYKLGIKEQNIIKQIILSKDNHPLSIEYNEEKNSIILAINNIYKEDF 833
Db 771 -LSPTE-----KEIMKRRDTP-----BFD- 788
QY 834 KCFTSGGDPNLPFYVTFNMLKLIKEKGNLTLYVPSAIWNNESSRILRKHIFARY---KLN 890
Db 789 -----IPVTFIVHSSKLLKNEGVLGFIIPSSFGTGVRYSNLRKELFTYKMLKLI 838
QY 891 Y-----IYQFNKKRKFQVHVSFKFAIFQLSNIKE 920
Db 839 YLPEDVFGAYVDCNIIILHKPKPKSEDLVLIYAPKTK----- 878
QY 921 STSSFKAFEMIQSSDNILKEITRDLKDSKDAY-KGIELNINOIKLSPIQESIIEPKDN 979
Db 879 -----KISPEFKNDLFIEYSKILNDPKCRIFPKSPIIYILDKIKONCRESLYLEDL 931
QY 980 BEFTLINMFKSF---SALGEGYIDFKG-----LPSIKNRKSLKCECNKNLIPLY 1029
Db 932 TEST-IGILASKYKESDKKENEYIPLYEGNVRYETKLNKYVDVFSKHKNKELINLF 990
QY 1030 SGANIHQFNSRFFEDKAKESSKLLWIDKEDLEKVLTKDNQYQTERVFVRAIASNTNERT 1089
Db 991 -----MSPEKIFIRIV-NRQDRI 1008
QY 1090 MISTLSPGNCYCVNS-----IYNDEKTPISLYKLFPIISFNSFVDFLRRFVDSN--V 1143
Db 1009 MASY---GNIEGWKKDLVVFVKPDPINYP---YLLGIILNSELISYI---YIGKSAIA 1059
QY 1144 LKSCLYQCPMPQPEKEILSNPLYNLAKNTSLLIAKNDPENFKYLLYLEYFKDKKVN 1203
Db 1060 LKDDFRQTTLELRELPIV-----IPKQMIINA-----LTQLSKURFELND-- 1101
QY 1204 KILKLDKEDEFFKEKENENNFIISLYSLAKEDFITLLGDFKALKNNKKGGEDIYISSLING 1263
Db 1102 ---KLNENDRIPL-----NIIDSLVYGIYFQDLIP-----KEELNEICNE--INGIICK 1146
QY 1264 YD 1265
Db 1147 YD 1148

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:30:59 ; Search time 78.7244 Seconds
(without alignments)
1341.095 Million cell updates/sec

Title: US-10-688-058-16
Perfect score: 6538
Sequence: 1 MKTNDIVKTNPNISLYKQL.....SSLIKGYDNYLLNNKIFYHK 1277

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/6_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/H_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/RE_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	394	6.0	1218	2	US-09-605-703B-600
2	382	5.8	1009	2	US-09-605-703B-600
3	316	4.8	1602	2	US-09-605-703B-600
4	316	4.8	1621	2	US-09-605-703B-600
5	316	4.8	1639	2	US-09-605-703B-600
6	315	4.8	2184	2	US-09-605-703B-600
7	309	4.7	622	2	US-09-605-703B-600
8	285.5	4.4	2662	2	US-09-605-703B-600
9	285.5	4.4	2663	2	US-09-605-703B-600
10	280	4.3	990	1	US-08-957-005-9
11	280	4.3	990	1	US-08-957-005-9
12	278	4.3	990	1	US-08-957-005-9
13	276.5	4.2	1007	2	US-08-957-005-9
14	269	4.1	2710	1	US-08-405-496A-6
15	269	4.1	2710	1	US-08-405-496A-6
16	269	4.1	2710	2	US-08-915-136-6
17	269	4.1	2710	2	US-08-915-136-6
18	269	4.1	2710	2	US-08-915-136-6
19	269	4.1	2710	2	US-08-915-136-6
20	265.5	4.1	1457	2	US-09-134-000C-6740
21	262.5	4.0	2777	2	US-09-543-681A-6124
22	261.5	4.0	1786	2	US-08-973-462-8
23	259	4.0	1068	2	US-09-966-997-11
24	255	3.9	1477	2	US-09-830-230A-414
25	255	3.9	1494	2	US-09-830-230A-413
26	252	3.9	3169	2	US-09-453-702B-257
27	252	3.9	3169	2	US-10-114-170-257

28	250	3.8	1119	2	US-09-830-230A-321	Sequence 321, App
29	247	3.8	1381	2	US-09-662-254B-20	Sequence 20, Appl
30	246	3.8	1087	2	US-09-830-230A-322	Sequence 322, App
31	245	3.7	3878	2	US-09-914-259-11	Sequence 11, Appli
32	239	3.7	648	2	US-10-338-898-2	Sequence 2, Appli
33	238	3.6	905	2	US-09-248-796A-16333	Sequence 16333, A
34	237	3.6	1173	2	US-09-248-796A-19313	Sequence 19313, A
35	236.5	3.6	1290	2	US-10-360-101-220	Sequence 220, App
36	236.5	3.6	2733	2	US-09-949-016-11433	Sequence 11433, A
37	236.5	3.6	3259	2	US-09-949-016-6507	Sequence 6507, Ap
38	235.5	3.6	3135	1	US-08-323-170B-2	Sequence 2, Appli
39	235.5	3.6	3135	2	US-08-954-441-2	Sequence 2, Appli
40	233.5	3.6	618	2	US-09-299-378-4	Sequence 4, Appli
41	233.5	3.6	1010	2	US-09-134-001C-5178	Sequence 5178, Ap
42	233	3.6	1257	2	US-09-245-928A-18	Sequence 18, Appli
43	232	3.5	1257	2	US-08-947-823-5	Sequence 5, Appli
44	229.5	3.5	10182	2	US-09-134-001C-3159	Sequence 3159, Ap
45	227	3.5	1078	2	US-09-248-796A-20284	Sequence 20284, A

ALIGNMENTS

RESULT 1

US-09-605-703B-600
; Sequence 600, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 600
; LENGTH: 1218
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-605-703B-600

Query Match 6.0%; Score 394; DB 2; Length 1218;
Best Local Similarity 19.0%; Pred. No. 6e-17;
Matches 211; Conservative 189; Mismatches 379; Indels 332; Gaps 42;
QY 331 YKTIITFNLLDK-----GSDAIKFPVNGGLFSDKVKYLNNEGLLSI 374
DB 27 YDSLQLLFTVLDKGHHPTDAKAVAFDAEAGEGLHFNLSADLFLPAATELIDRVG-LSN 85
QY 375 SIEELIVKMLFEEKNIKDFVKYSLDPSKFGELYETLLLEYDLRIA-----DTTVHR 429
DB 86 EALNKVLENLLSRVOSGKDRGFISVATLGVTELGOVLEGLMSYTGFIQAQEDLFEVAPHG 145
QY 430 IIEDGVYL-----IRTEELENKKVVKI-ATYLNKGNLYLTSRSLDRKKSQA 474
DB 146 KADKGSWMLPVSKADEVPADSFIEVDQAPGGGVKVRKRHRPGSVFVRQSRDRERSAS 205
QY 475 YTPDOLTFMVISSIEQLTKY-----SPIDKIINDNSCGSGHFLISCLDYLTKEVWY 528
DB 206 FYTPQVLTSTFTQATEELQASKRITTANDVLSLTICEPAMGSGAPAVEAVRLAE-LVL 264
QY 529 ELDKEDVKKELDEEYRVIIIESEEDVDQDSISKEL-VLKRMLLKCCIYGVNDINLSVEI 587
DB 265 EL-ROEELEQQIPAEADR-----AKELQKVKAHIALHQVYGVVDINLSTAVEL 308

QY 947 DSKDD--AYKGIELNINOIKKLSPIQESIIIEPKDNEFTLINOMFKSPFSGALGEGYIDPKK 1004
Db 670 DSBENTCVSSBKYNSDPWIFLSPETEAV-----FTKFTB--AQFEKLGE-ITDISV 718
QY 1005 GL-----DPSIKNRKSLKE-----C-----NNKGLI 1026
Db 719 GLQTSADKIYIFIPENETSDTYIFNYKGRYEIEKSIKCCPAIYDLSFGSPESIQNAQMI 778
QY 1027 FLYSGANIHPGRFPEDKDAKSSKLLWIDKEDLEKVLTKDNOYOTERVFYRAIASNTN 1086
Db 779 FPY---EIRDEEAYLLEETLENDYPLAWNYLNEFEKLEKRSIQGRNPKWYQVGRSQSL 835
QY 1087 ER-----TWISLSPGNCYCVNSIYINDEKTPISLYKPLFIISIFNSFVDFELRR 1137
Db 836 SKFHDKEKLIWTIATKPPVVLDRNLLFTGGNGP---YGLINOSIYLSLHYF----- 886
QY 1138 FVDSNVLSKLYQCPMPQPEKEILSNPLYNLAKNTSLIIAKNDPENFKYLLYLEYFKF 1197
Db 887 -----LGLSHPVIESWVKARA-----SEFRG-----SYSH 913
QY 1198 DKEKVNKI-----LKLKDEBFFK-----EK-----ENENN----- 1223
Db 914 GKQFIEKIPIRKIDFDDQDVEDVKYNTVTTVEKLIITTDRIKSESNGPRRRMLRRILDAL 973
QY 1224 -----FIASLYSLAKEDFITILGD 1243
Db 974 SNOLIQVINELYNDEEYTTVLND 998

RESULT 3
US-09-269-874A-7
; Sequence 7, Application US/09269874A
; Patent No. 6933130
; GENERAL INFORMATION:
; APPLICANT: Bujard, Hermann
; TITLE OF INVENTION: Recombinant Process for Preparing a
; TITLE OF INVENTION: Complete Malaria Antigen, GP190/MSPI
; FILE REFERENCE: GRUE-003
; CURRENT APPLICATION NUMBER: US/09/269, 874A
; CURRENT FILING DATE: 1999-08-02
; PRIOR APPLICATION NUMBER: PCT/EP97/05441
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: DE 19640817.2
; PRIOR FILING DATE: 1996-10-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1602
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-09-269-874A-7

Query Match 4.8%; Score 316; DB 2; Length 1602;
Best Local Similarity 19.7%; Pred. No. 9e-12;
Matches 295; Conservative 229; Mismatches 481; Indels 490; Gaps 73;

QY 2 KTNDI-----VKTNNPNIISLYKQLSKDFIK-----KENINKLDPFILIKNKLF 45
Db 172 KLVDCANDYCOIPFNKLIRANELDVLLKLVFGYKPLDNKONVGMEDY--IKNK-- 227
QY 46 SIDDNSTEANIESLLKYIFELNYSVEQQAGQIGVESRVDILLFENDKDKASFNKVK 105
Db 228 -----KTNIENELI-----EESKTDKNKNATKEBEKK-----LYQAQYDLSIYNKQLE 274
QY 106 EA-----KKNNEPIEDIL-IIAEVR-----PTFSFDAKDKLES 141
Db 275 EAHNLISVLEKRIIDTLKKNEN-----IKELLDKINEIKNPPANSNGTPTNLLDKNKKIEH 331
QY 142 E-----DQLYR-----YLNQYOKHYGILS-----NGKVMR 166
Db 332 EKXIKETAKTIKFNIDSLFTDPLEYLLREKKNIDISAKVTKSTEPNEVPNGVTYP 391
QY 167 L-YDKSKVLYGKRYI-----EFNFXKIEBEYKEQWFLVFIYLIRKERYLKTNSVIE 220

Db 392 LSYNDINNALNELNSFGDLINPPDYTKPSKNIYTDNE-----RKKFINEIKEKIK 442
QY 221 VEKEQIAKEKEIITQ---KTLKEILYERPPDSDIVFKIAKNIYDKEF-----KLSGK 267
Db 443 IEKKKIESDKSYEDRSKSLNDITKEYE-----KLLNEIYDSKFNNNDILTNFKWMGK 496
QY 268 -----EITOHILASILEESIIFILRIP-FIAYIED-----NDIFKKILOENKLYRSSIS 315
Db 497 RYSYKVEKLTHTNTFASYENSKHNLEKLTAKALYMEDYSLRNTIVVEKELKYKLNLSKIE 556
QY 316 FRYFFVDENTKK---KLEYKKLIITIENLDDKG---SDAIKPPVFNGLFSEDKVKYLYNN 368
Db 557 NEIETIVENIKDEEQOLFKEKTIKDNKPKDEKILEVSDIVKVQV----- 600
QY 369 EGLLSISEIEELVAKMLFFEEKNIKDEKFKYSRLDPKSGELYETLLEYD---LRIADT 425
Db 601 QKVLLANKIDELKKTQLILKNVELKHNHVPNSYKQENKQEPYLYLVKKEIDKLVFMP 660
QY 426 TVHRIITDGVYLRTEEBEELKKNVKIATYLVKNIVLTSRSLDRKXSGAYTTDPDDLTFDM 485
Db 661 KVESLINEEKNIKITEGQSDNSEPS-----TEGEITQATTKPGQOAGSALEGDS--- 710
QY 486 VISSIEEOLKTKSPLDIKI-----LDNSCGSGHPLISCLDYLTEKVM----- 527
Db 711 VOQAQEQKQAOFPVPVPEAKAQVTPPPAPVNNKTEN---VSKLDYL-EKLYEFLNT 765
QY 528 -----YELDKFEDVKKELDEYRVIIEESEYDVQDSISKELVLRMLLKXCIYGVGIN 581
Db 766 SYICHKYILVSHSTWNEKILKQYKITKEESKLSLSCDPLD-----LLFNIIQ-N 812
QY 582 PISVEITMLSLMINTPIFGTPLS--FIEHHIKAGNALLGYTKDEFFDIVKKKESGESL 639
Db 813 NIPVMYSMPDLSNNS-----LSQLFWEIYERKEMVNCVLYKLDN--DKIKNLLEEA--- 860
QY 640 KKEIKELIITLEDIYQKIK-----GINDTKE-----DIEKSKKIY 675
Db 861 KKVSTSVKTLSSSSMQPLSLTPQDKPEVSANDTSHSTNLNLSKLUFENILSGKNKIY 920
QY 676 KEYEESKIDNLN--RIIFSLIKLYLSFKSLNMFSDIASVISLIENTILGNKTSSEDK- 732
Db 921 QELIGQSSSENFYEKILKDSDDTFYNESFTNFVKSADDINSL-----NDESKKKL 971
QY 733 -EKIEKIRKLSYYKFFHYGIBFPDIOEGFDIVIGNPPWEKTKFNTEFESKH--IPNYR 789
Db 972 EEDINKLKK-----TLQLSFDL-----YNYKYLKLERLFDKKTGVKY- 1009
QY 790 KLGIKEONILKQELSKDNHPLSIEYNEEKNISIIANNIYKPFKCFCTSGDNLFRYFV 849
Db 1010 KMQIKKLTLLKEQLESKLN-----SLNN----- 1032
QY 850 TPNLKLKEKGNLTLYLPSAINWESSRILRKHIFARYKLYNIYQFENKKRFDKDVHSSPK 909
Db 1033 -----PKVLQNFVS-----FNKKKEAEIABT-- 1055
QY 910 FAIFQLSNIKESTSPFAKFMIQSSDNILKEITRLDKSDKDDAYKGIELNIN-QIKKLS 968
Db 1056 -----ENT-----LENTKILKCH-----YKGLVKNYNGESSPLKT 1085
QY 969 IQESIIEFKDN-----BEFTLINOMFKSPFSGALGEGYIDFKK-----GLDPSIKRKS 1016
Db 1086 LSEESIQTEDNVASLENFVLSKLEGL-----KDNLNLEKKLSYLSGSLHLHLIABLK 1141
QY 1017 LKECNKNKILFLYSGANIHQFNRSRFPEDKDAKSSKLLWIDKEDLEKVLTKD---NOY 1071
Db 1142 IK---NKN-----YTGNSPSENNT---DVNNALYESKFFLPEDTDVATVVSSESQSDTLEQS 1191
QY 1072 QTERVFYRAIASNTNERTMISTLSPGNCYCVNSIYINDEKTPISLYKPLFIISIFNSFV 1131
Db 1192 QPKKPASTHVGAESNIT-----TSQNVDDDEVDV-----IIVIFGSESE 1232
QY 1132 DF-----LLRRFVDSNVLK-----SCLYQCPMP---QPEKEILSNPLYNLAK 1172

Db 1233 DYDDLGOVVTGEAVTPSVIDNLSKIENEYEVLYLKLPLAGVYRSLLKQLLENNVMTFNVV 1292
Qy 1173 NTSLLAKNDPENFKVLLYLEYKFP-DKEKVNKILKLDKEDEF-FKEKENENNFIASLY 1230
Db 1293 KDILNSRFNKRNFKNVLESDLIPYKDLTSSNVVVK-----DPYKFLNKEKRDKFL--SSY 1346
Qy 1231 SLAKEDFIT-----LLGDFKALKNNKKE-----DYISSLIKGYDNYL--LNN 1271
Db 1347 NYIKSDITDINFANDVLGYKILSEKYSKDLSDLSIKKYINDKQGENEKYLPFLNN 1401
RESULT 4
US-09-269-874A-5
; Sequence 5, Application US/09269874A
; Patent No. 6933130
; GENERAL INFORMATION:
; APPLICANT: Bujard, Hermann
; TITLE OF INVENTION: Recombinant Process for Preparing a
; TITLE OF INVENTION: Complete Malaria Antigen, GP190/MSPI
; FILE REFERENCE: GRUE-003
; CURRENT APPLICATION NUMBER: US/09/269,874A
; PRIOR FILING DATE: 1999-08-02
; PRIOR APPLICATION NUMBER: PCT/EP97/05441
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: DE 19640817.2
; PRIOR FILING DATE: 1996-10-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1621
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-09-269-874A-5

Query Match 4.8%; Score 316; DB 2; Length 1621;

Best Local Similarity 19.7%; Pred. No. 9.1e-12;

Matches 295; Conservative 229; Mismatches 481; Indels 490; Gaps 73;

Qy 2 KTDNI-----VKTNNPISLYKQSKDPIK-----KENINKLDPFILIKNKLP 45
Db 191 KLVNDVANDYQIPFNLKIRANELDVLKLVFGYRKPLDNIKNVGRVEDY--IKNK-- 246
Qy 46 SIDDNSTANESLLKYPFELNYSVQKAGQLEGVESVDIILFENDKDKASFNK 105
Db 247 -----KTNIENELI-----ESKKTIDKNKATKEBEKK-----LYQAQYDLSLYNKQLE 293
Qy 106 EA-----KKNNEPIPIEDIL-IIAEVVR-----PTFSFDKADKLKES 141
Db 294 EAHNLISVLEKRIIDLTKNEN--IKELLDKINEIKNPPANSNGTPTNLLDKNKKIEEH 350
Qy 142 E-----DQLYR-----YLNQYQKHGILS-----NGKWR 166
Db 351 EKEIKEIAKTIKFNIDSIFTDPLEYLYLREKKNIDISAKVETKESTEPNEYPNGVTYP 410
Qy 167 L-YDKSKVLYGEKRYI-----EFNFXTKEEKEEYKEQEWFLFYILIRKERYLKTNSVIE 220
Db 411 LSYNDINNALNELNSFGDLINPFDYTKEPSKNITDNE-----RKKFINEIKKIK 461
Qy 221 VEKEQIAKEKEIIQ---KTLKEIILYRPDDSVIFKIAKNIYDKEF-----KL-SGK 267
Db 462 IEKKIESDKSYEDRSKSLNDITKEYE-----KLLNEIYDSKFNNDNIDLTNFEKMGWK 515
Qy 268 -----EITQHILASILESGIIFILRIP-PIAYED-----NDIFPKILQNKLYRSIS 315
Db 516 RYSYKVEKLTHTNTFASYSKHNLEKLTALKATKMTMEDYSLRNIIVVEKELKYQNLISKIE 575
Qy 316 FRYFFYDENTKK--KLEYKIIITFNLLDKG-----SDAIKFPVFNGLPSEDKVKYLLN 368
Db 576 NEIETLVENIKDEQLFEKKTIDENKPDKEKILEVSDIVKQV----- 619
Qy 369 EGLLSISIEBILVKMLPFEEKNINDEKFKVYSLRDPKPSFGELYETLLEYD---LRIADT 425
Db 620 QKVLLMKNIDELKKTQLTLKVNKLKHNHVNYSYKQENKQRPYLLIVLKBIDKLVKVP 679

Qy 426 TVHRIIEGCVYLIRTEEBLENKKNKIATYKGNLYLTLSRSLDRKKSQAYTTPDOLDTDM 485
Db 680 KYESLINEKNIKTEGQSDNSEPS-----TGEITGQATTKPGQAGSALSGLSDS----- 729
Qy 486 VISSIEBQKTSKPSDIDI-----IDNSCGSGHFLISCLDYLTBKVM----- 527
Db 730 VQAQAEQKQAPPPVPVPEAKAQPVPAPVNNKTEN-----VSKLDYL-EKLYEFLNT 784
Qy 528 -----YELDKFEDVKKDELDEBYRVIIESEEDYVDQDSISKELVLKMLKXCIYGVGIN 581
Db 785 SYICHKYLIVSHSTWNEKILQYKITKBESKSLSCDPLD-----LLFNIQ-N 831
Qy 582 RISVEITMLSLWINTFIFGTPLS--FIEHHIKAGNALLGYTKDEFFDVIKKKPFESGSLF 639
Db 832 NIPVMSWFDLSNNS-----LSQLFMEIYEKEMVCNLYKLKON--DKIKNLLSEA--- 879
Qy 640 KKRKEIITIILBDIYQIK-----GINDTTKE-----DIEKSKKIY 675
Db 880 KKVSTSVKTLSSSMQPLSLTPQDKPEVSANDTSHSTNLNLSKLFPENILSLGKNKIY 939
Qy 676 KEYEESKDIDL--RIIFSLIKLYSLSPKSLNMEFSDIASVISLIENILGNKTSSEDK- 732
Db 940 QELIGQKSSNFYEKILKDSDTFYNESFTNFVKSADINDSL-----NDESRRKKL 990
Qy 733 -EKIEKIRKLSVYKFFHYGIEFPDIOEGFDIVIGNPPWEKTKFNETEFPSKH--IPNYR 789
Db 991 EEDINKLKK-----TLQLSFDL-----YNYKYLKLERLFDKKTKVGY- 1028
Qy 790 KLGIKEQNIIOBILSKONHPLSIEYNEEKNSIIAINNIYKDFKCFSTSGDNLFRYFV 849
Db 1029 KMQIKKLTLLKEQLBSKLN-----SLNN----- 1051
Qy 850 TFWNLKLIKGNLYLVPSAINWESSRILRKHIFARYKLNLYIYQFENKPKVDHSSFX 909
Db 1052 -----PKHVLQNFV-----PFNKKEABIAET-- 1074
Qy 910 FAIFQLSNIKESTSSPKAFMIQSSDNILKEITRDLKDSKDAYKGIELNIN--QIKKLS 968
Db 1075 -----ENT-----LENTKILKH-----YKGLVKYNGESSPLKT 1104
Qy 969 IOESIIEPKDN-----BEFTLINRKFSGALGEGYIDFKK-----GLDPSIKNRKSL 1016
Db 1105 LSEESIOTEDNYASLENFKVLSKLEGL-----KONLNLEKKKLSYLSGLHLIAELKEV 1160
Qy 1017 LKECNNKILFYSGANIHOFSRPFEDKAKESSKLLWIDKEDLEKVLTKD-----NOY 1071
Db 1161 IK---NKN-----YTGNSPSENNT--DVNNALSEYKFKFLPEGTDVATVVSSEGSTLEQS 1210
Qy 1072 QTERVFYRALASNTNERTMTLSLPGNCYCVNSIYINDEKTPISLYKCLPIIISFNSFVF 1131
Db 1211 QPKKPASTHVGAESNIT-----TSQNVDDDEVVV-----IIVIFGSEEE 1251
Qy 1132 DF-----LLRRFVDSNVLK-----SCLYQCPMP---QPEKEILSNPLYLNLAK 1172
Db 1252 DYDDLGOVVTGEAVTPSVIDNLSKIENEYEVLYLKLPLAGVYRSLLKQLLENNVMTFNVV 1311
Qy 1173 NTSLLAKNDPENFKVLLYLEYKFP-DKEKVNKILKLDKEDEF-FKEKENENNFIASLY 1230
Db 1312 KDILNSRFNKRNFKNVLESDLIPYKDLTSSNVVVK-----DPYKFLNKEKRDKFL--SSY 1365
Qy 1231 SLAKEDFIT-----LLGDFKALKNNKKE-----DYISSLIKGYDNYL--LNN 1271
Db 1366 NYIKSDITDINFANDVLGYKILSEKYSKDLSDLSIKKYINDKQGENEKYLPFLNN 1420

RESULT 5

US-09-269-874A-3

; Sequence 3, Application US/09269874A

; Patent No. 6933130

; GENERAL INFORMATION:

; APPLICANT: Bujard, Hermann

; TITLE OF INVENTION: Recombinant Process for Preparing a

; TITLE OF INVENTION: Complete Malaria Antigen, GP190/MSPI
 ; FILE REFERENCE: GRUE-003
 ; CURRENT APPLICATION NUMBER: US/09/269,874A
 ; PRIOR FILING DATE: 1999-08-02
 ; PRIOR APPLICATION NUMBER: PCT/EP97/05441
 ; PRIOR FILING DATE: 1997-10-02
 ; PRIOR APPLICATION NUMBER: DE 19640817.2
 ; PRIOR FILING DATE: 1996-10-02
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 1639
 ; TYPE: PRT
 ; ORGANISM: Plasmodium falciparum
 ; US-09-269-874A-3

Query Match 4.8%; Score 316; DB 2; Length 1639;
 Best Local Similarity 19.7%; Pred. No. 9.2e-12;
 Matches 295; Conservative 229; Mismatches 481; Indels 490; Gaps 73;

QY 2 KTDNI-----VKTNNPISLYKQSDPIK-----KENINKLQDFILIRKPLF 45
 DB 191 KLDVDCANDYCOIPFNLKIRANELDVLKLVFGYRKPFLDNKDNVGMEDY--IKKKN-- 246
 QY 46 SIDDNTEANIESLLKYIFELNYSVEQKAGQIEGVESRDVILLFENDKDKASFNKXLK 105
 DB 247 -----KTINENELI-----EESKTDKNKNATKEBEKK-----LYQAQYDLSIYNKQLE 293
 QY 106 EA-----KKNNEPIEDIL-IAEVRK-----PTPSFDAKDKLKE 141
 DB 294 EAHNLISVLEKRDITLKNEN--IKELLDKINEKNPPANSNTPTLLDKNKIEEH 350
 QY 142 E-----DQLYR-----YLNQYQKHGILS-----NGKWR 166
 DB 351 EKEIKETAKTIKENIDSLFDPLELYLREKKNIDISAKVETKESTEPNEPVGWYTP 410
 QY 167 L-YDKSKVLYGKRYI-----ENFXKIEEKEYKQEWFLYILIRKERYLKTNSVIE 220
 DB 411 LSYNDINNALNELNSFGDLNPFDTYKPSKNITONE-----RKKFINEIKERIK 461
 QY 221 VEKEQIAKEEIIQ---KTLKEILYERPDSDSIVFKIAKNYDKEF-----KLSGK 267
 DB 462 IEKKIESDKSVEDRSKSLNDITKEYE-----KLLNEIYDSKFNNDILDTWFERQMGK 515
 QY 268 -----EITQHILASILEESIIFILRIF-FIAYIED-----NDIFKKILOENKLYRSSIS 315
 DB 516 RYSYKVEKLTHNTFASVENSKENLEKLTALXKYMEDYSLRNIVVEKELAYKYNLISKIE 575
 QY 316 FRIFFYDENTK---KLEYKIITIFMLDKG-----SDAIKPPVFNGLFSEDKVKYVLNN 368
 DB 576 NEIETLVENIKDBEQLEFKKITDKENKPKDEKILEVSDIVKVQV----- 619
 QY 369 EGLLSISEIEILVKMLFPEEKIKDKFVKYKRLDPKSGELAYETILEYD---LRIADT 425
 DB 620 QKVLMMNKIDELKKTQLILKNVLEKNHVPNSYKQENKQEPYLVILVRKEIDKLKVFMP 679
 QY 426 TVHRIEDGVYLIRTEBELENKVNKIATYLGKNIYLTSLDRKSKSGAYTTPDDLTDFM 485
 DB 680 KVESLINEEKNIKTEGQSDNSES-----TEGEITQATTKGQQAAGSALEGDS----- 729
 QY 486 VISSIEQLTKSPLDIKI-----IDNSCGSHFLISCLDYLTEKVM----- 527
 DB 730 VQAQAQEQQAQPPVPVPEAKAQPPTPAPVNNKTEN---VSKLDYL-EKLYEPLNT 784
 QY 528 -----YELDKPFDVKELDEEVRVIERSEEDVDQDSISKELVLKMLLXCIYGVGIN 581
 DB 785 SYICHKYLVSHTMNEKILQYKIKTEEESKLSKSDPLD-----LLFNIIQ-N 831
 QY 582 PISVEITMLSLWNTFTFGTPLS--PTEHHIKAGNALLGYTKDBFFDIVKKKESGSLF 639
 DB 832 NIPWYMSFSLNNS-----LSQLFWEIYEKEMVCNLYKLKN--DKIKNLLIEA----- 879
 QY 640 KGRKEIITILEDIYQKIK-----GINDTTKE-----DIEKSKKIY 675

DB 880 KKVTSVVKILSSSQMPLSLTTPQDKREVSANDTSHSTNLSNLSKLPEMILSGKNKIY 939
 QY 676 KEYBESKDNL--RIIFSLIKLYSLSPDKSLNMFSDIASVISLIENILGNKTSBEDK- 732
 DB 940 QELIGQKSSSENFVEKILKSDTFYNESFTNFVKSADDINSL-----NDESKRKKL 990
 QY 733 -EKIEKIRKLSYKFFHYGIEFPDIOEGFDIVGNPPWEKTKFNETEPESKH--TPNYR 789
 DB 991 EEDINKLKK-----TLQLSPDL-----YNNYKYLKLERLFDKKTGVGY- 1028
 QY 790 KLGIKEONIIKOBILSKDNHPLSIEYNEEKNSIIAINIYKDFPKCFTSGDGNLFRYFV 849
 DB 1029 KWOIKKLTLLKEQLESKN-----SLNN----- 1051
 QY 850 TPNLKLKEKGNLTLYLPSAINWESSRILRKHIFARYKLNLIYIYQENKGRFQDVHSSF 909
 DB 1052 -----PKHVLQNFSV-----FFNKKEABEIAET-- 1074
 QY 910 FAIFQLSNIKESTSSPKAFMIQSSNIIKEITRDLKSDKDAYKGIELNIN-OIKKLSP 968
 DB 1075 -----ENT-----LNTKILKKH-----YKGLVYKYNESGSPKLT 1104
 QY 969 IOESIIEPKDN---EEFTLINKMPSKFSALGRGYIDFKK-----GLDPSIKNRKSL 1016
 DB 1105 LSEESIQTEDNVSASLENFKVLSKLEKL-----KDNLEKKKLSYSSGLHHLIAELKEV 1160
 QY 1017 LKECNKNKILFYSGANIHQFNSRFPEDKAKESSKLLWIDKEDLEKVLTKD-----NOY 1071
 DB 1161 IK---NKN---YTGNSPSENNT---DVNNALESYKKFPEGTDTVATVVSSESGDTLEQS 1210
 QY 1072 QTERVEFYRAISNTNERTWISTLSPGNCYCVNSIYINDSKTPISLYKULFIISFNSFVF 1131
 DB 1211 QPKKPASTHVGAESNTIT-----TSQNVDDDEVDDV-----IIVPIFGESEE 1251
 QY 1132 DP-----LLRRFVDSNVLK-----SCLYQCPMP---QPEBEKILSNPLYNLAK 1172
 DB 1252 DYDDLGQVVTGAVTSPSIDNLSKIENIEYEDVLYLPLAGVYKSLKKQLENNVMTFNVV 1311
 QY 1173 NTSLLIAKNDPBNFKYLLYLEYKFK-DKEKVNKILKLDKEDEF-FKEKENENNFIASLY 1230
 DB 1312 KOILNSRFKNENFKNVLESLLIPYKDLTSSNVVVK---DPYKFLNKEKRDKFL--SSY 1365
 QY 1231 SLAKEDFIT-----LLGDFKALKNKKGE-----DYSSLIKGVNDYL--LNN 1271
 DB 1366 NYIKDSITDINFANDVLGYKILSEKYSKSLDSIKKYINDKQGENEKYLPFLNN 1420

RESULT 6
 US-09-417-485D-6
 ; Sequence 6, Application US/09417485D
 ; Patent No. 6541202
 ; GENERAL INFORMATION:
 ; APPLICANT: Long, David M.
 ; APPLICANT: Metz, Anneke M.
 ; APPLICANT: Love, Ruschelle A.
 ; TITLE OF INVENTION: Telomerase Reverse Transcriptase (TERT) Genes
 ; FILE REFERENCE: 47714-5009-US
 ; CURRENT APPLICATION NUMBER: US/09/417,485D
 ; CURRENT FILING DATE: 2002-06-14
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 2184
 ; TYPE: PRT
 ; ORGANISM: Plasmodium falciparum
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (330) (335)
 ; OTHER INFORMATION: Xaa at position 330 = Leu or Ile;
 ; OTHER INFORMATION: Xaa at position 335 = Asp or Gly.
 ; US-09-417-485D-6

[illegible]

RESIN.T 8

```

US-09-595-684B-31
; Sequence 31, Application US/09595684B
; Patent No. 6544766
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Onashi, Cara
; APPLICANT: Sakowicz, Roman
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Yu, Ming
; TITLE OF INVENTION: Human kinesins and methods of producing
; FILE OF INVENTION: and purifying human kinesins
; FILE REFERENCE: cytop036
; CURRENT APPLICATION NUMBER: US/09/595,684B
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 2662
; TYPE: PRT
; ORGANISM: Human
; US-09-595-684B-31

Query Match 4.4; Score 285.5; DB 2; Length 2662;
Best Local Similarity 20.08; Pred. No. 1.6e-09;
Matches 292; Conservative 222; Mismatches 452; Indels 493; Gaps

QY 1 MKTNDIVKTNW-----PNTSLYKQLSKDFFIKGNINKL-----KD 35

```

Db 2145 KEEQHECINKFEMDFIDEVEKQKELLIKIQLHQQDCDVPSELRD-----LKLNQ 2194
Qy 963 IKKLSPTQSIIEFKDNEFTLNKMFSEFSGALGEGVIDFKGLDPSIKNRKSLKCECN 1022
Db 2195 NMDLH-IEELLKDFSS-EPPSIKTEFQV-----LSNRKEM-----2229
Qy 1023 KNLIFLYSGANIHQFNSRFFEDKAKESKLLMI-----DKEDLEKVLTKDQ--YQTERV 1076
Db 2230 -----TQFLEE-----WLNTRFDIEKLKNGIKQENDRICQVNNF 2263
Qy 1077 FYRAIASNTNERTMTSLPGNCYCNSIYINDEKTPISYKGLFIISFNSFVDFLLR 1136
Db 2264 FNNRIIAINNESTEFERSA-----TISKWEQDLKSLKEKNEKL-----2303
Qy 1137 RFVDSNVLSCLYQCPMPQPEEKEILSNPLYLNKNTSLIAK-NDPENFKVLLYLEVP 1195
Db 2304 -FKNYQTLKTSLASGAQVNPFTQD-KNPNHVTSGA--TQTTTEKIRELENSLH-----EAK 2355
Qy 1196 KPDKEKVNKILKDKDEFFPEKEKENENNFIIASLYSLAKE-----DFITLLGDFFKA 1246
Db 2356 ESAMHESKIIKQKELEV-----TNDIIAKLQAKVHESNKCLEKTKETIQVLQDKVA 2408
Qy 1247 L---KNKKGEDYISSLIK 1262
Db 2409 LGAPKYKEETEDLKMVLK 2427

RESULT 9

US-09-538-092-1252
; Sequence 1252, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapSeqFormatter Version 0.9
; SEQ ID NO 1252
; LENGTH: 2663
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number Q02224
US-09-538-092-1252

Query Match 4.4%; Score 285.5; DB 2; Length 2663;
Best Local Similarity 20.08; Pred. No. 1.6e-09;
Matches 292; Conservative 222; Mismatches 452; Indels 493; Gaps 72;

Qy 1 MKTNDIVKTN-----PNISLYKQLSKDFIKKENINKL-----KD 35
Db 1266 INTQLEKSHTKLQEEIPVLHEEQELLPNV---KKVSE---TQETWNELELLTEQSTTKD 1319
Qy 36 PFILIKNKLPSIDN-----STEANIESI-----LKYIFELNYSVEQQAGQIEGVSER 85
Db 1320 STTLARIEMERLRLNEKFOESQEEISLTKERDNLTKIKEALEYKHDQLK-----EHIRET 1375
Qy 86 VDILLFENDKDKASFNKLEAKGNPEPIEDILIIAEVKRPFFSDAKD-----136
Db 1376 LAKIQESQSQEQSLNKK-----EKDNETTK-----IVSEMEQ-----FKPKDSALLRIEIE 1422
Qy 137 -----KLKESDQLYRLNQYQKHGYILSGKWRILYDKSKVLYGKRYIEFNFXXKIE 190
Db 1423 MGLSKRLQSHDSEKSVAKEDD-----LQRLQEVLSQSDQLENIKEIVA 1470

Qy 191 KEEYKEQEWFLVFIYLIRKRYLKTSNVIEVEKEQIAKEKEIIQKTLKEILYERPDSDIV 250
Db 1471 KHELETEBELKVAHCCCKEQE---ETINELRVNLSKEETEISTIQOLEAI-----NDKLO 1522
Qy 251 FKIANIYDKPKLGGKITQHILASILEESIIFILRFFIAYIEDNDIFPKILQBNKLY 310
Db 1523 NKI-QBIYEKBPQNIKQISE-----VQENV-----NEL-KQPKHEKAK 1560
Qy 311 RSISFRPFYDENTKKLEYKIIITIPNLLDKGSDAKFPVPFNGGLFSDVKVYLNNEG 370
Db 1561 DSAL-----QSIE-SKMLELTNRLQESQEEIQIMI-----KKEEMKRVOR 1600
Qy 371 LLSI-----SBIELVQMLFFEKNIKDEKFKVYKRLDPKSFGLYELTLEYDLRIAD 424
Db 1601 ALQIREDQLKENTKEIVAKM---KESQEKYQFLKWTAVN-----ETQEKMEIE-----1647
Qy 425 TTVHRIEDGVYLRTEBELENKNVKNKIATYILKGNIIYLS-----RSLDRKSGAY 475
Db 1648 -----HLKEQFETQKLN-LENISTENIRLTQILHENLEEMRSVTTKER-----1688
Qy 476 YTPDDLTDPMVLISSIEEQKTKSPDLIDIKIIDNSCGSGHFLISCLDYLTKVWVLDKFPD 535
Db 1689 ---DDL-----RSVEETLKVERD---OLKEN-----LRETITRDLEKQEE 1722
Qy 536 VK-----KELDEEYRVIIIE-----SEEDVDQD-SISKELVLKRL 570
Db 1723 LKIVHMLKHEQETIDKLGIVSEKTNESNNQKOLEHNSDALKAQDLKIQBELRAHMH 1782
Qy 571 LKXCIYQVD-INPISVEITMLSLWINTFIFGTPLSFIEHHIKAGNALLGYTKDEFFDIVK 629
Db 1783 LKEQOETIDKLGIVSEKT-----DKLSNMQKDLNSNAKLEQIQL-----1825
Qy 630 KFPESGSLFKKRIKXI-----ITILEDIYQIKIGINDTK-----EDIEKSKKIYKEVESK 682
Db 1826 KANEHQLIITLKKDVNSETQKVSEMEQLKKQIKQDSLTKSLKTIENLJLAQELHENLEEMK 1885
Qy 683 DI-----DNLRIIFSLIKL-----YSLSPDKSLNMEFSDIASVISILNIGLTKTSSDK 732
Db 1886 SVMKERDNRVETUKEERDQKESLOETKARDLEIQOELKTARML-----SKEHK 1937
Qy 733 EKIEKIRKLSYYKFFHYGIEFPDIQEGFD-----IIVGNPPWE 771
Db 1938 ETVDKLE-----KISEKTIQISQDKDLKSKDELQKQIQLKQKELQLLRVKEDVMS 1992
Qy 772 KTKFNETEFPKXHI-PNYRKLGIKEQNIQKQILSKONHPLSIEYNE-----EKN 820
Db 1993 HKKINEMEQLKKQFEPNY-----LCKCEM---DNFQLTKKJHESLEIRIVAKERD 2040
Qy 821 SIITAINNIYKDFKCFKFTSGDPNLPFRYFTFNLKLIKKEGN-----LTVLP 867
Db 2041 ELURIKESLKNERD-----QFIATIREMIARDQNHQVKEPKRLSDGQOHLME 2089
Qy 868 SAINNESSRILRKHIPARYK-----LNYI-YQFENKCRFKDVHSSPKFAIFOLSN 918
Db 2090 SL--REKCSRI--KELLKRYSEMDDHVECLNRLSLDLEKEIEFHRIMKKLKYVLSYVTKI 2145
Qy 919 KESTSSFKAKFMIOSSDNILKE-----ITRDLKSDKDAYKIELNINQ 962
Db 2146 KEEQHECINKFEMDFIDEVEKQKELLIKIQLHQQDCDVPSELRD-----LKLNQ 2195
Qy 963 IKKLSPTQSIIEFKDNEFTLNKMFSEFSGALGEGVIDFKGLDPSIKNRKSLKCECN 1022
Db 2196 NMDLH-IEELLKDFSS-EPPSIKTEFQV-----LSNRKEM-----2230
Qy 1023 KNLIFLYSGANIHQFNSRFFEDKAKESKLLMI-----DKEDLEKVLTKDQ--YQTERV 1076
Db 2231 -----TQFLEE-----WLNTRFDIEKLKNGIKQENDRICQVNNF 2264
Qy 1077 FYRAIASNTNERTMTSLPGNCYCNSIYINDEKTPISYKGLFIISFNSFVDFLLR 1136
Db 2265 FNNRIIAINNESTEFERSA-----TISKWEQDLKSLKEKNEKL-----2304

QY 1137 RFVDSNVLKSLYQCPMPQPEKEILSNPLYNLAKNTSLIAK-NDPENFKYLLYLEYF 1195
 Db 2305 -FNYQTLKTSLASGQVNTTQD-KNPNHVTGRA--TQUTTEKIRELENSLH-----EAK 2356
 QY 1196 KFDKRVNKLKLDKEDDFPKKEKENENNFIASLYSLAKE-----DFITLLGDFA 1246
 Db 2357 ESAMHESKIIKQKELEV-----TNDIIAKLOAKVHESKCKLEKTETIQVLQDKVA 2409
 QY 1247 L-----KNKKGEDYISSLIK 1262
 Db 2410 LGAKPYKEETEDLMKLVK 2428

RESULT 10

US-08-392-625-20
 ; Sequence 20, Application US/08392625
 ; Patent No. 5837485
 ; GENERAL INFORMATION:
 ; APPLICANT: Entian, Karl-Dieter
 ; APPLICANT: G tz, Friedrich
 ; APPLICANT: Schnell, No. 5837485bert
 ; APPLICANT: Augustin, Johannes
 ; APPLICANT: Engelke, Germar
 ; APPLICANT: Rosenstein, Ralf
 ; APPLICANT: Kaletta, Cortina
 ; APPLICANT: Klein, Cora
 ; APPLICANT: Wieland, Bernd
 ; APPLICANT: Kupke, Thomas
 ; APPLICANT: Jung, G nther
 ; APPLICANT: Kellner, Roland
 ; TITLE OF INVENTION: Biosynthetic Process For The Preparation
 ; TITLE OF INVENTION: Of Chemical Compounds
 ; NUMBER OF SEQUENCES: 42
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
 ; STREET: 1100 New York Avenue
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/392,625
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/876,791
 ; FILING DATE: 30-APR-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Esmond, Robert W.
 ; REGISTRATION NUMBER: 32,893
 ; REFERENCE/DOCKET NUMBER: 0652.0980002
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 371-2600
 ; TELEFAX: (202) 371-2540
 ; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 990 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear

US-08-392-625-20
 Query Match 4.3%; Score 280; DB 1; Length 990;
 Best Local Similarity 20.0%; Pred. No. 1e-09;
 Matches 243; Conservative 186; Mismatches 394; Indels 390; Gaps 64;
 QY 6 IVKTNPNISLYKQSKDFKKNINKLKDFLLKKNLFP-----SID- 48
 Db 16 MVRTPIFSIELYNQ-----FLKSDNI-----DYDLILQNDIFPKESIMTTTNYQSIGKIDW 67

QY 49 --DNSTEANI--ESLLKYIFBELLNYSVEQKAGQIGVESRVDILLPENDKDKASFNKLK 105
 Db 68 EKDNKTRNVKESLLKYL--IRMSTRSTPYGMLSGV-----ALG 104
 QY 106 EAKKNNEPIPIEDILLIAEVRGPTFSFADKDKLKESEDOLYRYLNOYQKHGILSNGKW 165
 Db 105 EFSENN--IKID-----SSFHKD--VKIDGOWLYKLVHLYESDY----- 142
 QY 166 RLYDKSVLYGKRYIEFNFXKIEEKESYKQEWFLFYILIRKERYLKTNSVIEVEKEQ 225
 Db 143 TYKDSFVIWNOQNYIYNNRLYLDNNSITEN-----KENDVLSVKYNS 186
 QY 226 IAKEKEIIQTKLEILYERPDSDIVFKIKNYIDYKFKLSGKEITHOILASILEEIIIFI 285
 Db 187 IL-----VFTHENSKNITVEELVQLISSKYSIENKEEVFVQOELINKELIIF- 234
 QY 286 LRIFFIAYIEDNDIFKKILOENKLYRSSISPRFFYVDENTKKLEYKIIITENLADKGS 345
 Db 235 -----SDLRPTLENKPL-----DYIINSLAPKNSL-----VOTLININEIT 272
 QY 346 DAIKFPVNGGLFSEDKVYLNNEGILLSISEIEELVQMLFFBEKNIKDEKFKVYKR--- 402
 Db 273 KYSKMPLGKG-----EYKYLDIVNLMS-----QLFVSKNYLQIDTYIDYSRNL 316
 QY 403 -----LDPKSPGELYETLLEYDLRIADTTVHRIIEDGYLIRTEEELE 445
 Db 317 KQSLADNISEAAAYILWLLSPHEFGT--KTIRNY-----HEFFMDKYGF-----EQLV 361
 QY 446 NKK-----VNKIATYLGNIYLTSLD-----RKKSQAYYTPDLDLTF---MWI 487
 Db 362 NLKQLSLDINGFG--YPKDSYSFSNNIAFLKELYLAIQNNSHIEITENDVKNLEKNNTV 420
 QY 488 SSIEEQLKT-----KSPLDIKIIDSCGS-----GHFLISCLDYLTE 524
 Db 421 SKINAPVSTEIYSEIYFGNSIKGYEDFAVISPLGSGFNAGATGRTGNF----- 470
 QY 525 KVMYELDKFEDVKELDBEYRVIIIESEYDVODSISKELVLK-----RMLLKCIYV--- 577
 Db 471 ----NIKKKNQLQKEIVHHYNNYMNEN--DLEISQLNEAPLNSRVNVLNNRNYTCL 523
 QY 578 -----VDINPISVEITMLSLWNTFFIGTPLSFIEHHIKAGNALLGYTKDEFDVK 629
 Db 524 NLNLPKSDIDINDIFIGATFNKLYL-----YSEKH-----DSRIVFVSNMFN--- 566
 QY 630 KXFESGSLFKKRIKEIITILEDI--YOKIKGINDTTKEDIEK-----SKYIKY--- 678
 Db 567 --YFSGELYK-----FLREISEFKTFQITEEGIDSLPFCPRIYKNILKPAT 616
 QY 679 -----EESKDIDNLRIFSLIKYLSLSPDKSLNMFSDIASVLSLIENILGNKTSSEDK 733
 Db 617 WKINSEMFSETENLNRFPATIRKWHI--PKDVIIFAGD-----NRLLLNLDKHLILKK 670
 QY 734 KIEK---IRKLSYKFFHYGIEFPDIQSGFDIVIGNPPWEKTKFNETSEFF-----SKHI 785
 Db 671 ELKCHGRIRILSEF-----INESNNRMLIEV--TPLYKKTSLKEQSFITPKNRKHF 721
 QY 786 PNVR-----KLGIK---EQNIKQETLSKDNHPLSTEYNEKNSIATINN--YKPDFKCF 836
 Db 722 NNLKDFSIHLSPKTYQDNFIQDYL-----PFITE-----LKVNFINKFYIKF 768
 QY 837 TSGGDPNLPFRYFVTFNLKLIKKEGNLYLVPSPAINWNESSRILRKHIFARYKLYN---IY 893
 Db 769 KEDED-----FI--KLRLREDEDSQIV-SFIKNWKDYCLNSELSDYSIDVYVPEV 819
 QY 894 QENKGRFKDVHSSKFP-AIFOLSNIKESTSFKAFKMIQSSDNL-----KEITRDLKD-- 947
 Db 820 RYGGPHVIDIENFNFYDLSLDSIIQSEFKIPKEFIVAISIDFLDLVLEINKSEKEEI 879
 QY 948 ---SKODAYKGLN--INQIKKL-SPIQESI---EPKDNBEFTLINKWFSKFSALGEG 998
 Db 880 LINNAEDLYRSNDIREYKNLLAK/TNPKNDYELKKEFPNLHEF-----LFNKISILE-- 932

QY 999 YDFFKGLDPSIKRKSLLKCCNNKLNILFLYSGANIHQFNSRFFEDKDKAKSSKLLWIDK 1058
 Db 933 --NLKXTLQSLVTSRSRI-----IGSFHMRNCNRIFGINPEKEFVL----- 973
 QY 1059 EDLEKVLTKDQY 1071
 Db 974 -SIFNEITTKY 985

RESULT 11
 US-08-466-961A-20
 ; Sequence 20, Application US/08466961A
 ; Patent No. 5843709
 ; GENERAL INFORMATION:
 ; APPLICANT: Entian, Karl-Dieter
 ; APPLICANT: G tz, Friedrich
 ; APPLICANT: Schnell, No. 5843709bert
 ; APPLICANT: Augustin, Johannes
 ; APPLICANT: Engelke, Germar
 ; APPLICANT: Rosenstein, Ralf
 ; APPLICANT: Kaletta, Cortina
 ; APPLICANT: Klein, Cora
 ; APPLICANT: Wieland, Bernd
 ; APPLICANT: Kupke, Thomas
 ; APPLICANT: Jung, G nther
 ; APPLICANT: Kellner, Roland
 ; TITLE OF INVENTION: Biosynthetic Process for the Preparation of
 ; TITLE OF INVENTION: Chemical Compounds
 ; NUMBER OF SEQUENCES: 42
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Sterne, Kessler, Goldstein & Fox
 ; STREET: 1100 New York Avenue, NW
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/466,961A
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/392,625
 ; FILING DATE: 22-FEB-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/876,791
 ; FILING DATE: 30-APR-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/784,234
 ; FILING DATE: 31-OCT-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Esmond, Robert W.
 ; REGISTRATION NUMBER: 32,893
 ; REFERENCE/DOCKET NUMBER: 0652.0980004
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 371-2600
 ; TELEFAX: (202) 371-2540
 ; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 990 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; US-08-466-961A-20

Query Match 4.3%; Score 280; DB 1; Length 990;
 Best Local Similarity 20.0%; Pred. No. 1e-09;
 Matches 243; Conservative 186; Mismatches 394; Indels 390; Gaps 64;
 QY 6 IVKTNPNISLYKLSKDFIKENINKLKOFFILIKNLF-----SID- 48

Db 16 MVRTPIFSIELYNQ-----FLKSDNI-----DYLLILQNDIFKESIMTTTNYLYQSIGKIDW 67
 QY 49 --DNSTANI--ESLLKYIFEEELNYVEQKQKAGIEGVSERVILLFENDKDKASFNNK 105
 Db 68 EKDNKTRNVKESLLKYL---IRMSTRPTYNLSGV-----ALG 104
 QY 106 EAKKNEPIPIBDILIIAEVKKRPTFSFAKDKLKESEDOLYRYLNQYQKHGILSNKVM 165
 Db 105 EFSENN-IKID-----SSFHKKD-VKIDGOWLYKLVHVESDY----- 142
 QY 166 RLYDKSVLYGKRYIEFNFXKIEBEKYEQEWFLVLYIRKERYLTSTNVIEWEKEQ 225
 Db 143 TYKDSFVIWNOQNYNNRLYLDNNSITEN-----RRNDVLSVKYNS 186
 QY 226 IAKEKEIIQKTLKEILYERPDOSIVFKIAKNYDKEFKLSGKEITQHILASILEESIIFI 285
 Db 187 IL-----VFIHNSKKNITVEELVQIISKYSIENKEEVKFVQELINKELIIF- 234
 QY 286 LRIFPIAYIEDNDIFPKILQENKLYRSSISPRYFFYDENTKKKLEYKIIITIFNLDDKGS 345
 Db 235 -----SDLRPTLENKNPL-----DYIINSLNPKNSL-----VGTLINISNEIT 272
 QY 346 DAIKFPVFNGLFSEDVKVYLNNEGLLSISEIEILVKMLFFEEKNIKDEKPFVKYSR--- 402
 Db 273 KYSKMPLGKG-----EYKYLDIVNLS-----QLFVSKNYLQIDTYIDYSRNEL 316
 QY 403 -----LDPKSFGELYETLLBYDLRIADTTVHRRIEDGVYLIIRTEEELE 445
 Db 317 KOSLADNISEAAAYILWLJSPHEFGT--KTIRNY-----HEFFMDKYGF-----EQLV 361
 QY 446 NKK-----VNKIATYKGNILYTSRSLD-----RKKSGAYVTPDLDLTF-----MWI 487
 Db 362 NLKQLLSDINGFG-YPKDSYSFSNNIAFLKKEYLLAIONNSHIEITENDVNLEKNVT 420
 QY 488 SSIEBOLKT-----KSPLDIKIIDSOGS-----GHPLISCLDYLTE 524
 Db 421 SKINAPVSTEIYSEIYFGNSIKGYEDFAVISPIILSGFNAGATFGRPTGNF----- 470
 QY 525 KVMYELDKFEDVKELDEEYRVIIIESEYDVQDSISKELVLK-----RMLLKCIYG--- 577
 Db 471 ----NIKKQNLOKEIVHHNNYNNMEN---DLEISQLEAPLNSRVNLIANNRIYNTCL 523
 QY 578 -----VDINPISVEITMLSLWINTFIQTPISFIEHRIKAGNALLGYTDEFFDIVK 629
 Db 524 NLNLPKSDIDINDIFIGATFNKLYL-----YSEKH-----DSRIYFVSNMFEN--- 566
 QY 630 KXFEGSFLPKRIKEIITILEDI-YOKIKGINDTTKEDIK-----SKKIYKY----- 678
 Db 567 --YFGSELYK-----FLREISPEKTKFIQITEEGIDSLPFCPRIIYKNILKPAT 616
 QY 679 -----EESKDIDNLRIFIISLILKLYSLFDSKLSNMEFSDIASVISLIENTLGNKTSSEDE 733
 Db 617 WKINSEMFSETENLNRNFATIRKWHI--PKDVIIFAGD-----NRLNLNDLKHILILKK 670
 QY 734 KIEK--IRKLSYKFFHYGIEFPDIQGFQIVIGNPPWEKTKFNETEFF-----SKHI 785
 Db 671 ELKKGRIILLESF-----INESNNRMLEIV--TPLYKKTSLKEQSIIIPKRNKHF 721
 QY 786 PNYR-----KLGIK---EQNIKQEILSKDNHPLSIYNEEKNSIINAIN-IVYFOPKCF 836
 Db 722 NNLKDWFSIHLSIPKTYQDNFIQDYLL-----PFITE-----LKVNNFINKFFYIKF 768
 QY 837 TSGGDPNLFYFVTNFKLKEKGNITLVPSAIWNSESSRIIRKHFARYKLVN---IY 893
 Db 769 KEDED-----FI--KLRLLEDEDSQIY-SFIKNWKDYCLLNSELSIYDIYVPEVY 819
 QY 894 OPENKKRKFQDVHSSKFF-AIFQLSNIKESTSSKAKFMIOSSDNIL---KEITRDLKD-- 947
 Db 820 RYCGPHVEDIENFFMYDSDLSINIIQSEFKIPKEFIVAISIDFLDLYLEINKSEKEI 879
 QY 948 ---SKDDAYKGIELN--INQIKKL-SPIOESI---EFKDNESFTLKNKMFSPKSALEG 998

Db 880 LINNAEDYSDIREYKNNLAKLTNPKNDEYELKKEFPNLFH-----LFNKISILE-- 932

Qy 999 YIDPKGLDPSIKRNRKSKLLKECNKNNKLI FLYSGANIHFQNSRFFEDKAKESKLLWIDK 1058

Db 933 --NLKKTLOKSLATSRRI-----IGSFHMRNCNRIFGINPEKEKFLV----- 973

Qy 1059 EDLEKVLTKDNOY 1071

Db 974 -SIFNEITTKYKY 985

RESULT 12

US-08-645-193B-15

Sequence 15, Application US/08645193B

Patent No. 5962253

GENERAL INFORMATION:

APPLICANT: Kupke, Thomas

APPLICANT: Gotz, Friedrich

APPLICANT: Kempter, Christoph

APPLICANT: Jung, Gunther

TITLE OF INVENTION: Oxidative Decarboxylation of Peptides

TITLE OF INVENTION: Catalyzed by Flavoprotein EpiD

NUMBER OF SEQUENCES: 70

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/645,193B

FILING DATE: 13-MAY-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Esmond, Robert W.

REGISTRATION NUMBER: 32,893

REFERENCE/DOCKET NUMBER: 0652.1540000

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 990 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: not relevant

MOLECULE TYPE: protein

US-08-645-193B-15

Query Match 4.3%; Score 278; DB 1; Length 990;

Best Local Similarity 20.1%; Pred. No. 1.3e-09;

Matches 243; Conservative 183; Mismatches 402; Indels 380; Gaps 63;

Qy 6 IVKTNNDNISLYQLSKDPTKGNINIKLKDFFILIRKFLP-----SID- 48

Db 16 MVRTPIFSIELYQ---FLKSDNI---DYDLILQNDIFKESIMTTTNYLYQSIGKIDW 67

Qy 49 --DNSTEANI-ESLLKYIFEELNYSVEQKAGQIEGVESRVDILLFENDKDKASFNKKL 105

Db 68 EKONKTRNKESLLKYL---IRNSTSTPYGMLSGV-----ALG 104

Qy 106 EAKKNEPIEDILIIAIEAKRPTFSFDAKDKLKESEDQLYRYLNQYQKHGILSNKQW 165

Db 105 EFSENN-IKKD-----SSPHKD-VKIDQWLYKLVLHLYESDY----- 142

Qy 166 RLYDKSKVLYGKRYIEFNFKIEEYKQBWFLFIYLIRKERYLTKNSVIEVEKEQ 225

RESULT 13

US-09-957-005-9

Db 143 TTYKDSFVWQNYIYNNRLYLDNNSITEN-----KENDVLSVKYNS 186

Qy 226 IAKEKEBIIOKTKEILYERDDSIIVFKIAKNIYVDKFKLSGKEITOHILASILEESIIFI 285

Db 187 IL-----VFTHENSKKNIITVEELVQLISSKYSTENKEEVVQELINKKEIIF- 234

Qy 286 LRIFFIAYIEDNDIFKKILQENKLYRSSISFRFFYDENTKKGLYKKIITINLLDKGS 345

Db 235 -----SDLRPTLENKNPL-----DYIINSLNPKNSL-----VGLTINISNEIT 272

Qy 346 DAIFPVFNGGLFSEDKVKYLNNEGLLSISEIEILVKMLFFBEKNIKDKFKVKYSR-- 402

Db 273 KYSKMPLGKG-----EYKYLDIVNLS-----QLFVSKNYLQIDTYIDYSNEL 316

Qy 403 -----LDPKSGELYETLEVDLRIADTTVHRIIEDGYVLIRTEELE 445

Db 317 KQSLADNISAAAYILMLSPNHFGT--KTIRNY-----HEFFMDKYGF--EQLV 361

Qy 446 NKK-----VNKIATYLGNIYLTSLD-----RKSGAYTTPDLDLTF--MWI 487

Db 362 NLQQLSDINGFG-YPKODSYSFNSNIAFLKEKYLLAIQNNSHIEITENDVKNLEKNNTV 420

Qy 488 SSIIEQLKT-----KSPLDIKIIDNSCGS-----GHFLISCLDYLTEKVMY 528

Db 421 SKINAPVSTIEIYSEIYFGNSIKGYEDPAVISPILGSNAGATGRTGN-----F 470

Qy 529 ELDPEDVKKELDEEYRVIIERSEYDVQDSISKELVLK-----RMLLKCIYG----- 577

Db 471 NIKKKQLOKEIVHHYNNYNNEN---DLEISQLNEAPLNSRVNVLNNRIYNTCLNML 527

Qy 578 -----VDINPISVEITMLSWINTFIRGTPLSFIHHIKAGNALLGTYKDEFFDIVKKKFE 633

Db 528 PKSDIDINDIFIGATFNKLYL-----YSEKH-----DSRIVFVSMSFN-----YE 568

Qy 634 SGFSLFKRIKEIITILEDI-YOKIKGINDTTKEDIEK-----SKTIYKEV----- 678

Db 569 FGSLEYK-----FLREISFEKTFIQPIIEEGIDSLPFCPRIYKNIILPATWKN 620

Qy 679 -EESKDIDNLRIFSLIKLYSLFDSKLNMEFSDIASVISLIENILGNKTSSEDKKIEK 737

Db 621 SEMFSETENLNRFAFIR-EKWHIPKDVIIAFGD---NRLNLLNLDKHLIILKELKK 675

Qy 738 ---IRKLSYYKFFHYGIEFPDIOEGFDIVIGNPPWEKTKFNTEFF-----SKHIPNYR 789

Db 676 HGRIRILESP-----INESNRMLEIV--TPLYKTSLSKQSFIIIPKNNRKHFNKLK 726

Qy 790 -----KLGIK---EQNIKQELISKDNHPLSIEYNEEKNSIIAINN-IYKFDKFCFTSGG 840

Db 727 DMFSIHLSIPKTYQDNFIQDYLL-----PFITE-----LKVNNPINKFFYIKFEDE 773

Qy 841 DPNLFRYFVTNKLKIKERGNLTLYVPSAIWNESSRILRKHFARYKLYN---IYOFEN 897

Db 774 D-----FI--KLRLREDEYQIY-SPIKWKDYCLNSELVDYSDIVDYVEVRYGG 824

Qy 898 KKRFDVHSSFKPAIFOLSNIKESTSSFKAKFIQSSDNIL---KEITRDLK-----SK 949

Db 825 PHVIEDIENFFMYDLSLSINIIQSEFKIPKEFIVASIDFLDYLEINKSEKEILINNA 884

Qy 950 DDAYKGIELN--INQIKKL-SPIQESII---EPKDNEEFTLNMKFSKLSALGEGYIDFK 1003

Db 885 EDLYRSNDIREYKNNLAKLTNPKNDEYELKKEFPNLFH-----LFNKISILE-- 935

Qy 1004 KGLDPSIKNRKSKLLKECNKNNKLI FLYSGANIHFQNSRFFEDKAKESKLLWIDKEDLEK 1063

Db 936 KTLQKSLYTSRSRI-----IGSFHMRNCNRIFGINPEKEKFLV-----SIFN 977

Qy 1064 VLTKNQY 1071

Db 978 EITKTKKY 985

Query Match 4.1%; Score 269; DB 1; Length 2710;
Best Local Similarity 19.4%; Pred. No. 1.9e-08;
Matches 304; Conservative 257; Mismatches 499; Indels 506; Gaps 82;

QY 1 MKTNDIVKNNPNISLYKQSKDFIKENINKLKDFILLKKNLKFIDNNTSTANIESLL 60
DB 492 LKASDLIEKFP-----ENLSQTEQEI---NSLWFDQASAKYQFE--- 531
QY 61 KYIFEELNVSQOQAGIEGVSVDIILLFEND--KDKASFNNKLEAKKNEPIED 118
DB 532 KYVRDYTGSSLSBD-----NGVDFNKVNTALDKNLLNKNIPSNVVEAGSKY- 579
QY 119 ILIIAEVKRPTSFDAKDKLSEDO---LVRYLNOYKHGYILSGKVRWLYDKSKVL 174
DB 580 VHYIQLQDDISYEATCNLFSPKNRNSIIQRNMNESAKSYFLSDG----- 627
QY 175 YGEKRYIBFNFKIEKEBYKEQWFLFYIIRKERYLKTSNVIEVEKEQIAKEBIIQ 234
DB 628 ---ESILELNKYRIPERLKNKEK---VKYTFIGHKDEFTWSEFARLSVDLSNEISSFL 681
QY 235 KTLKEILYERPDOSIVFKIAKNYDKFKLS-----QKEITQHILASILEE- 280
DB 682 DTRIK--LDISPKNVENVLLGCNMFSDVFNVEETYPGK-----LLLSIMDKITSTLPDVNK 734
QY 281 -SIIFILRIFFAYIEDNIFKILQENKLYRSSISFRYFFYDENTKKLEYKIIITFN 339
DB 735 NSIT-----IGANQYEVRAINSEGRKELLAHSGRWINKBEATMSLSKEYI--FFD 783
QY 340 LLDKGSDAIKFPVFNGLSPEDVKVLLNNEGLSISEIBELVKMLFEEKNIKDKBFVK 399
DB 784 SID-----NKLKAKSKNIPGLASISED-----IKTLLDASVPTKFI 823
QY 400 YS-RLD-PKSFGB-LYETLLEVDLRIADTVHRIIEDGVYLIRTEBEL-ENKKNVKNATY 455
DB 824 NNLKLNIESIGDYIYKLEPVKNIHNSIDDLIDFNLENVSDLEYELKKN----- 878
QY 456 LKGNILYTSRSLDRKSGAYTPDDLTDFWVSISSIEQKTKSPDLIKIDNSCGSHFL 515
DB 879 -----NLDEK-----YLISPEDISKNNSTYSVRFINSG----- 908
QY 516 ISCLDYLTAKVYVELDKFEDVKELDEYRVIIERSEYVDQDSISKEL-VLKBMLLKKC 574
DB 909 -----ESVVE-----TEKEIFSKY-----SEHITKEISTIKNSII----- 939
QY 575 IYGVINDIPASVEITML--SLWINTFIETGTPLSPIEHKIRAGNALLGYT--KDEFFDI--- 627
DB 940 ---TDVGNLLDNQLDHTSQVNTL---NAAPFIQ-----SLIDYSSNKVDLNDLSTS 986
QY 628 VK-----KKPESGF-----SLFKRKRIEIIITLEDIYOKI-----KGIN--DT 663
DB 987 VKVQLYAQLFSTGLANTYDSIQLVNLSNAVNDTINVLPITTEGIPVISTILDGINLGA 1046
QY 664 TKEDIKSKYIKEYESKDIDNRIIPSLIKLYSLSPDKSLNMEFS-----DIASVISLI 719
DB 1047 IKELDEHDPPLKKELEAK-VGVLAJNMSLSIAATVASIVGIGAEVTIFLLPPIAGISAGI 1105
QY 720 ENILGNKTSSEK--EKIEKIRKLSYKFFHYGIEFFDIOEGF-DIVIGNPWEKTFE- 775
DB 1106 PSLVNNELIHDKATSVNAYFNHLSKSKYGPPLKTEDDKILVPIDDLVISEDFNNNSIK 1165
QY 776 -----NETEFF-----SKHPN---YRKLGIKEQNI-IKQETLSKD 807
DB 1166 LGTCNILAMEGGSGHTVTGNDIDHFFSSPSISSHIPSLSIYSAIGIETNLDPSKIMWLP 1225
QY 808 NHPLSTIEY-----NEEKNSIITAINIY--KDFPKFTSGGDPNLPFRYFTV-- 850
DB 1226 NAPSFRVFWETGAVPGLRSLENDGTLLDSIRDLYPKFYWRFYA-----FFDYAITTL 1279
QY 851 -----FNLKLIKEGNLTLYVPSAIWNSSSRILKRIHAFARYKLVNIYQFNKKFKVD 904
DB 1280 KPVYEDTNIKIKDKDTRNFIPTITNE-----IRNKSYSFDGAG-GTYSLL 1327
QY 905 HSSFKFAIFQLSNIKESTSFKAQFMIOQSDNLIKETITRD---LKDSK---DDAYKGIELN 959

RESULT 15

US-08-405-496A-6

Sequence 6, Application US/08405496A

Patent No. 5919665

GENERAL INFORMATION:

APPLICANT: WILLIAMS, JAMES A.

TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM

TITLE OF INVENTION: NEUROTOXIN

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESSES: MEDLEN & CARROLL, LLP

STREET: 220 MONTGOMERY STREET, SUITE 2200

CITY: SAN FRANCISCO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/405,496A

FILING DATE: 16-MAR-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/329,154

FILING DATE: 25-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/161,907

FILING DATE: 02-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/985,321

FILING DATE: 04-DEC-1992
 PRIOR APPLICATION DATA: US 07/429,791
 FILING DATE: 31-OCT-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: INGOLIA, DIANE E.
 REGISTRATION NUMBER: 40,027
 REFERENCE/DOCKET NUMBER: OPD-01308
 TELEPHONE: (415) 705-8410
 TELEFAX: (415) 397-8338
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2710 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-405-496A-6

Query Match 4.1%; Score 269; DB 1; Length 2710;
 Best Local Similarity 19.4%; Pred. No. 1.9e-08;
 Matches 304; Conservative 257; Mismatches 499; Indels 506; Gaps 82;

1047 IKELLDEHDPKKELEAK-VGVLAInMSLSIAATVASIVGAEVTTIPLPIAGISAGI 1105
 QY 720 ENILGNKTSSEDK-EKIEKIRKLSYYKFFHYGFEFPIQESGF-DIVIGNPPWEKTKF- 775
 DB 1106 PSLVNNELILHDKATSVVYFNHLSSESKYGLPKTEDEKILVPIDDLISEIDFNNSIK 1165
 QY 776 -----NETEFF-----SKHIPN-----YRKLGIKEQNI-IKQILSKD 807
 DB 1166 LTCNILAMEGGSGHTVGNIDHFFSPSISSHIPSLSIYSAIGIETENLDSKIMMLP 1225
 QY 808 NHPLSIEY-----NEEKNSIANNIY--KPDFKCFSTSGGDNLFRIYVFT-- 850
 DB 1226 NAPSrvFWMETGAVPGLRSLNDGTRLLDSIRDLYPGKPYWRFYA-----FFDYAITLL 1279
 QY 851 -----FNLKLIKKEGNLTLYVPSAINWNESSRILRKHFARYKLYNIYQFENKKRFDV 904
 DB 1280 KPEVEDTNIKIKLDKOTRNFIMPTITNE-----IRKLUSYSDGAG-GTYSLL 1327
 QY 905 HSEKFAIFQLSNIKESTSSPKAKFMIOSSDNILKEITRD---LKDSK---DDAYKGIENL 959
 DB 1328 LSSYPIS-----TNINLS-----KDDLWIFNIDNEVEISEITNGIKKGLIKDVLSKIDIN 1379
 QY 960 INQIKKLSPIQESIIEFK---DNE-----EFTLINQKFKSPSALGEG 998
 DB 1380 KKKL-----IIGNQITIDFGSDIDNKDRIPLTCELDKISLIIEINLVAKSYSLISGDKN 1435
 QY 999 YIDFKKGLDPSIKNRKSLKECN-----NKNLIFLYSGANIHQFNRSFPFEDKAKESSKL 1053
 DB 1436 YL-----ISNLSNTIEKINTGLDSKNIAINYT-----DESNKKYF-GAISKTSQKS 1481
 QY 1054 LMKIDEDLEKVL-----TKDNOYQTERVYVR-----AIASNTNERTMIS 1092
 DB 1482 ITHYKDSKNILFEYNDSTLEFNSKDFIAEDINVPKMDINTITGKYVYDNTDKSIDFS 1541
 QY 1093 -TSPGNCYCVNSIYIN-----DEKTPISLYKKLFI--ISIFNSFVFD--- 1132
 DB 1542 ISLVSKNQVKVNGLYLINESVSYLDVFKNSDGHHTSNFMNLFNLDNISFWKLFGPENIN 1601
 QY 1133 FLRRF-----VDSN-----VLKSLYQC-----PMPQP 1156
 DB 1602 FVIDKYFTLVGKTNLGYVEFICDNNKNIDIFYGEMWTSKSTIPSGNGRNVVVEPIYNP 1661
 QY 1157 EEKEILS-----NPLY-----LNLAKNTSL-----LIAKN- 1181
 DB 1662 DTGEDISTLDYSYELGYIDRYINKVLAPDLYTSLINI--NTNYSNEYYPEIIVLNP 1719
 QY 1182 -----DPENPKY-----LLYLEYKPKDKEKVNKILK-----DKED 1212
 DB 1720 NTFHKVNINLSDSSPEYKMWSTEGSDFILVRYLEESNKKILQIRIKGIIISNTQSFNKM 1779
 QY 1213 EPFKE-KENENNPIIASLYSLAKEDFITL--LGDFKALKNKKKGEDYISLLIKGYDNYLL 1269
 DB 1780 IDPKIDKLSLGYIMSGNFKSFNSNELDRDLHG-FKIIDNKTYTYDEDSKLVKGLIN--I 1836
 QY 1270 NKKIFY 1275
 DB 1837 NNSLFY 1842

Search completed: January 24, 2006, 20:02:59
 Job time : 94.7244 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:57:15 ; Search time 224.44 Seconds
(without alignments)
2377.329 Million cell updates/sec

Title: US-10-688-058-16
Perfect score: 6538
Sequence: 1 MKTNDIVKTNPNISLYKQL.....SSLIKGNDYLLNNKIFVHK 1277

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA_Main:*
1: /cgn2_6/prodata1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/prodata1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6534	99.9	1277	5	US-10-688-058-16 Sequence 16, Appl
2	872	13.3	174	5	US-10-688-058-14 Sequence 14, Appl
3	455.5	7.0	1253	4	US-10-335-977-7399 Sequence 7399, Ap
4	435	6.7	1532	3	US-09-738-626-4321 Sequence 4321, Ap
5	419.5	6.4	1279	3	US-09-882-227-388 Sequence 388, App
6	392	5.8	1009	5	US-10-800-946-4 Sequence 4, Appli
7	354	5.4	1441	5	US-10-732-923-3352 Sequence 3352, Ap
8	337	5.2	2910	5	US-10-732-923-3342 Sequence 3342, Ap
9	335.5	5.1	1939	5	US-10-732-923-3340 Sequence 3340, Ap
10	328.5	5.0	1527	5	US-10-732-923-3354 Sequence 3354, Ap
11	324	5.0	1985	4	US-10-732-923-3351 Sequence 3351, Ap
12	318	4.9	740	4	US-10-335-977-7398 Sequence 7398, Ap
13	316	4.8	1602	5	US-10-354-924-7 Sequence 7, Appli
14	316	4.8	1621	5	US-10-954-924-5 Sequence 5, Appli
15	316	4.8	1639	4	US-10-087-464-10 Sequence 10, Appl
16	316	4.8	1639	5	US-10-954-924-3 Sequence 3, Appli
17	315	4.8	2184	4	US-10-304-095-6 Sequence 6, Appli
18	314	4.8	696	4	US-10-335-977-7397 Sequence 7397, Ap
19	307	4.7	1478	5	US-10-732-923-3353 Sequence 3353, Ap
20	301.5	4.6	1163	4	US-10-282-122A-51864 Sequence 51864, A
21	301.5	4.6	1163	5	US-10-732-923-3296 Sequence 3296, Ap
22	295.5	4.5	1875	4	US-10-369-493-22285 Sequence 22285, A
23	295.5	4.5	1875	5	US-10-732-923-3334 Sequence 3334, Ap
24	295.5	4.5	1875	5	US-10-732-923-3335 Sequence 3335, Ap
25	293	4.5	1178	4	US-10-282-122A-52434 Sequence 52434, A
26	293	4.5	1196	4	US-10-282-122A-52737 Sequence 52737, A
27	293	4.5	6761	5	US-10-732-923-15035 Sequence 15035, A

ALIGNMENTS

RESULT 1

US-10-688-058-16
; Sequence 16, Application US/10688058
; Publication NO. US20050058661A1
; GENERAL INFORMATION:
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: HALE, KATHERINE S.
; APPLICANT: JOHNSTON, STEPHEN A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF THE GENUS
; TITLE OF INVENTION: BORRELIA
; FILE REFERENCE: MCRO:003US
; CURRENT APPLICATION NUMBER: US/10/688,058
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 1277
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (186)
; OTHER INFORMATION: Xaa = anything
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (573)
; OTHER INFORMATION: Xaa = anything
US-10-688-058-16

Query Match 99.9%; Score 6534; DB 5; Length 1277;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MKTNDIVKTNPNISLYKOLSDFIKKENINKLKDFFILIKNKLPSIDDNSTEAIESLL	60
DB	1	MKTNDIVKTNPNISLYKOLSDFIKKENINKLKDFFILIKNKLPSIDDNSTEAIESLL	60
QY	61	KVIFELNYSVEQKAGQIEGVESRVDILLFENDKDKASFNNKLKEAKKNNEPIEDIL	120
DB	61	KVIFELNYSVEQKAGQIEGVESRVDILLFENDKDKASFNNKLKEAKKNNEPIEDIL	120
QY	121	IIAEVKRPTFSPADKDKLKESEDQLYRYLNQYKHYGILSNGKWLVDKSKVLYGEKRY	180
DB	121	IIAEVKRPTFSPADKDKLKESEDQLYRYLNQYKHYGILSNGKWLVDKSKVLYGEKRY	180
QY	181	IFBNFKIEKEEYKEQEWFLFYILIRKERYLKTNSVIEVEKEQIAKEKEIIQKTKEI	240
DB	181	IFBNFKIEKEEYKEQEWFLFYILIRKERYLKTNSVIEVEKEQIAKEKEIIQKTKEI	240

QY 241 LYERPDDSVFKIAKNYDKFVKLSGKEITQHTILASILEESIIFILRIFFIAYIEDNDIF 300
 DB 241 LYERPDDSVFKIAKNYDKFVKLSGKEITQHTILASILEESIIFILRIFFIAYIEDNDIF 300
 QY 301 KKILQENKLYRSSISPRYPFYDENTKKGLSYKKIITIFNLLDKGSDAIFPVFNGGLFSE 360
 DB 301 KKILQENKLYRSSISPRYPFYDENTKKGLSYKKIITIFNLLDKGSDAIFPVFNGGLFSE 360
 QY 361 DKVKYLNNEGLLSISIEEILVKMLFEEKNIKDEKFKVYSRLDPKSGELVETLLEYDL 420
 DB 361 DKVKYLNNEGLLSISIEEILVKMLFEEKNIKDEKFKVYSRLDPKSGELVETLLEYDL 420
 QY 421 RIADTTVTHRIIEOGVVLIRTEEBLENKVNKIATYLGNIYLTLSRLDRKKSAYTTPD 480
 DB 421 RIADTTVTHRIIEOGVVLIRTEEBLENKVNKIATYLGNIYLTLSRLDRKKSAYTTPD 480
 QY 481 LTDFMWISSIEEQLKTKSPLDIKI IDNSCGSGHFLISCLDYLTEKVMYELDKFEDVKEL 540
 DB 481 LTDFMWISSIEEQLKTKSPLDIKI IDNSCGSGHFLISCLDYLTEKVMYELDKFEDVKEL 540
 QY 541 DEYRVIIERSEBYDVQDSISKELVLRMLLKXCIYGVNDINPISVEITMLSLWINTFI 600
 DB 541 DEYRVIIERSEBYDVQDSISKELVLRMLLKXCIYGVNDINPISVEITMLSLWINTFI 600
 QY 601 TPLSFTEHHIKAGNALIGYTKDBFFDIVKKPFSGFLPKRIKEIITILEDIYQIKGI 660
 DB 601 TPLSFTEHHIKAGNALIGYTKDBFFDIVKKPFSGFLPKRIKEIITILEDIYQIKGI 660
 QY 661 NDITTKEDIEKSKIKYEBERSKDIDNLRIIFSLIKLYSLFSDKSLNMFSDIASVLSLIE 720
 DB 661 NDITTKEDIEKSKIKYEBERSKDIDNLRIIFSLIKLYSLFSDKSLNMFSDIASVLSLIE 720
 QY 721 NILGNTSSBDKIEKIRKLSYKPFHYGIFPDIOEQFDIVGNPPWEKTKFNETEF 780
 DB 721 NILGNTSSBDKIEKIRKLSYKPFHYGIFPDIOEQFDIVGNPPWEKTKFNETEF 780
 QY 781 FSKHIPNRYKLGITKEONI KOEILSKDNHPLSIYNEEKNSIITAINNIYKDFPKCFTSG 840
 DB 781 FSKHIPNRYKLGITKEONI KOEILSKDNHPLSIYNEEKNSIITAINNIYKDFPKCFTSG 840
 QY 841 DPNLFRYFVTFNLKIKEKGNLTLYVPSAIWNNESSRILRKHIFARYKLYNYIYQENK 900
 DB 841 DPNLFRYFVTFNLKIKEKGNLTLYVPSAIWNNESSRILRKHIFARYKLYNYIYQENK 900
 QY 901 FKDVHSSFKFAIFOLSNIKESTSFKAKFMIQSSDNILKEITRDLKSDKDAYKGIELNI 960
 DB 901 FKDVHSSFKFAIFOLSNIKESTSFKAKFMIQSSDNILKEITRDLKSDKDAYKGIELNI 960
 QY 961 NQIKLSPIOESIIEPKDNEEFTLINQWFKFSALGEGYIDFKGLDPSIKNRKSLKEC 1020
 DB 961 NQIKLSPIOESIIEPKDNEEFTLINQWFKFSALGEGYIDFKGLDPSIKNRKSLKEC 1020
 QY 1021 NNKNLIIFLYSGANIHQNSRFFEDKDAKSSKLMWIDKEDLEKVLTKDNQYQTERVY 1080
 DB 1021 NNKNLIIFLYSGANIHQNSRFFEDKDAKSSKLMWIDKEDLEKVLTKDNQYQTERVY 1080
 QY 1081 IASNTNERTWISTLSPGNCVNSIYINDEKTPISLYKKLPITISFNSFVDFLLRRFVD 1140
 DB 1081 IASNTNERTWISTLSPGNCVNSIYINDEKTPISLYKKLPITISFNSFVDFLLRRFVD 1140
 QY 1141 SNVLKSCLYOCMPQPEKEITLSNPLYNLAKNTSLIAIKNDPENPKYLLYLEYKFDKE 1200
 DB 1141 SNVLKSCLYOCMPQPEKEITLSNPLYNLAKNTSLIAIKNDPENPKYLLYLEYKFDKE 1200
 QY 1201 KVKKILKLDKDEDEFKEKENENFI IASLYSLAKEDFITLGDPKALKNKKKGEDYISSL 1260
 DB 1201 KVKKILKLDKDEDEFKEKENENFI IASLYSLAKEDFITLGDPKALKNKKKGEDYISSL 1260
 QY 1261 IKGYDNYLLNNKIFYHK 1277
 DB 1261 IKGYDNYLLNNKIFYHK 1277

RESULT 2
 US-10-688-058-14
 ; Sequence 14, Application US/10688058
 ; Publication No. US20050058661A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SYKES, KATHRYN F.
 ; APPLICANT: HALE, KATHERINE S.
 ; APPLICANT: JOHNSTON, STEPHEN A.
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
 ; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF THE GENUS
 ; TITLE OF INVENTION: BORRELIA
 ; FILE REFERENCE: MCRO:003US
 ; CURRENT APPLICATION NUMBER: US/10/688,058
 ; CURRENT FILING DATE: 2003-10-17
 ; NUMBER OF SEQ ID NOS: 141
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 14
 ; LENGTH: 174
 ; TYPE: PRT
 ; ORGANISM: Borrelia burgdorferi
 US-10-688-058-14
 Query Match 13.3%; Score 872; DB 5; Length 174;
 Best Local Similarity 99.4%; Pred. No. 6e-42;
 Matches 173; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 268 EYQHILASILEESIIFILRIFFIAYIEDNDIFKKILOENKLYRSSISPRYPFYDENTKK 327
 DB 1 QYQHILASILEESIIFILRIFFIAYIEDNDIFKKILOENKLYRSSISPRYPFYDENTKK 60
 QY 328 KLEYKKIITIFNLLDKGSDAIFPVFNGGLFSEDVKYLNNEGLLSISIEEILVKMLFF 387
 DB 61 KLEYKKIITIFNLLDKGSDAIFPVFNGGLFSEDVKYLNNEGLLSISIEEILVKMLFF 120
 QY 388 EEKNKDEKFKVYSRLDPKSGELVETLLEYDLRIADTTVTHRIIEOGVVLIRTE 441
 DB 121 EEKNKDEKFKVYSRLDPKSGELVETLLEYDLRIADTTVTHRIIEOGVVLIRTE 174
 RESULT 3
 US-10-335-977-7399
 ; Sequence 7399, Application US/10335977
 ; Publication No. US20040052799A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DOUGLAS SMITH et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
 ; RELATING TO HELICOBACTER PYLORI FOR
 ; DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 10031
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 28 State Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109-1875
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD/ROM ISO9660
 ; COMPUTER: IBM PC Compatible
 ; OPERATING SYSTEM: Windows NT 4.0
 ; SOFTWARE: UNIX
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/335,977
 ; FILING DATE: 30-Dec-2002
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/993,002
 ; FILING DATE: 17-DEC-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mandragouras, Amy E.
 ; REGISTRATION NUMBER: 36,207
 ; REFERENCE/DOCKET NUMBER: GTN-018
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 227-7400

```

; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 7399:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1253 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   HYPOTHETICAL: YES
;   ORIGINAL SOURCE:
;     ORGANISM: Helicobacter pylori
;   FEATURE:
;     NAME/KEY: misc feature
;     LOCATION: (B) LOCATION 1...1253
;     SEQUENCE DESCRIPTION: SEQ ID NO: 7399:
US-10-335-977-7399

Query Match      7.0%; Score 455.5; DB 4; Length 1253;
Best Local Similarity 20.7%; Pred. No. 3.4e-17;
Matches 302; Conservative 202; Mismatches 499; Indels 453; Gaps 61;

QY 22 KQPIKEN-----INKLKDFFILINKLPSIDDNSTEANIESLLK--YIFEELNYS 70
DQ 10 KVFIKKYNQPKETIENFEKINSLEENVPQDDDEEFQKNEINKFLKNAYGDCNTY- 68
QY 71 VEOQKAGQIEGVESVDILLFENDKDKASFNNKLKAKNNNEP-----IEDI- 119
DQ 69 -----EKVDYIYVDKEVRVLI-----EVKALNNKTEFPKRENPPLSKAFQOMVLYFLEEIE 120
QY 120 -----LIAEVKRPFPSPDAKDLKESD-OLYRVLYNOYKHGILSNGKWRLVD 169
DQ 121 KEKNSLKHTIICNA-HEFLFDCKDLSLKEDKRIKKFYKVAKEGT-----D 169
QY 170 KSKVLY-----GKRYIEFNFXKIEEKEQEWFVLIYILIRKERYLKTS 216
DQ 170 SSKPKFYKLEQFLOQSDFOGELRYTHFNL-----SDDFKS---LPLIYQVLSQEVLLKQK 221
QY 217 NVIEVEKEQIAKEKEIIQKTLKELLYERPDD--SIYFKIA--KXIYDKFELSGKEITH 272
DQ 222 RTLDN-----TLNKQFYELLYTLGLEEQNDKGLIKRSPRTKNSISGALKKEKYKNLDD 277
QY 273 ILASILEESIFILRIFFIAYIEDNDIFPKILQENKLYRSSISPRYF-----FYDENTK 326
DQ 278 EYVALL---IANNRILFLRLESLLISFEHFERSFLTTGNFKDFNALTFFEVLLAK 334
QY 327 KCLEYKIIITIFNLDKGSDAIFPVPNGGLFSE-----DKVKYLNNEGLSISEIEE 379
DQ 335 KNSRLPEIKENKILG-----KIPYLNSSLFDKTPLELKKGHEIKLLDNKKL-----E 381
QY 380 ILVKMLFFBEKNIKDKFKVYSRLDPKSPGELYETLLEVDLRAD-----TTVHRIIEDGV 435
DQ 382 IYKNSVLKKHENYQDK-----ASLPLLEFFFKPLHLYKFTTTPKDKDNT 427
QY 436 YLIRTEBELENKVNKIATYLNKNIYLTSLDRKSGAYTTPDDLTDFWVSSIE--- 491
DQ 428 ---DTSERLNPVS-----LGLVFEKNGYKEGSPFTSPFITSYCKESITIVIL 475
QY 492 -----EOLKT-----KSPLDIKIDNCGSGHGLFSLCLDYL 522
DQ 476 DKFNQKNTECDKLEKNVKNLSYKEDKRYQLQLLLTLRICDPVAGSGHFLVLSALNEM 535
QY 523 TEKVWYELDKFEDVKKLEDEYRVIIIESE-----FYDVQDSISKE 563
DQ 536 V-LIAYELGLIASLHR-----YSLTLENDEIITHYPTGTGIBFNFKKPLRENDPHHIIQKE 589
QY 564 LV-LKRMMLKXCICYGVNDINPISVEITWLSWI-----NTPIFGTPLSFIEH 608
DQ 590 LPELKKSIIENCFLGVNDINPNSCEITKRLWIELLYKYSYVIFBEGKNTNALET-LPNIDI 648
QY 609 HIKAGNALLGYTKDEFFDIVKKKFSFSLFKKRIKEIITIL-----EDLY 654
DQ 649 NIKCANSLI---SRFNLDLKK-----IPNKKIKQYKDLVAQYKDPNPLYNKADLI 701
QY 655 QKIKGINDT-----TKEDIEKS-KKIYKEYEESKQIDNLRIRIIFSLIKLYSLSFDKS 704

```

```

Db 702 NKIQDKNTFFSLTKDPTKKAELKTIKHIKYN-----FFAL-----DDKS 744
QY 705 LNMEFSDIASVISLIENILGN-KTSSDKE-----KIEKIRKL-----SSYKFFH 749
DQ 745 L-----LDGLNYFIPSLFGTLKSPKEEAEAFASYGRIRALRKLDLDDALSQGEYQNAFE 798
QY 750 YGIEFPDIQE-----GFDIVIGNPWEKTKENETFEFFSKHPIYKRLGKIQKNIKOEI 803
DQ 799 WRFEFPEVLDDRGDFLGFDCIIGNPPY-----IRQEQ 830
QY 804 LSKDNHP-LSIEYNEEKNSIIAINNIIYKFDKCFCTSGGDPNLFYFVTNFKLKEKGNL 862
DQ 831 I-KDLKPLLEKQYQDFYNS-----TADIYTFPALSFHLREKGS 870
QY 863 TYLVPSAIWNSSRILRKHIFARYKLYNYIYOFENKGRFQVHSGFKFAIFQLSNKEST 922
DQ 871 AFITSNKYARAYGAKLREWLKTKTLVSYMELNALK-----VFESAAVDTSI 918
QY 923 SSFKAKFMIQSDNLIKELITROLKDS-KDAYKGIELNINQIKLSPIQESIEFKDNEE 981
DQ 919 IHFIKQPLKESDFKYBPTDDKDDLKSTPYLSMKQNALSTESFIFANATLLDLRD--- 975
QY 982 FTLINKMFSKPSALGEGYIDFKGLDPSIK-----NRKSLKKEC-----NN 1022
DQ 976 -----KBSVGTPLKMDIQINYGIKGTGANEAFLIPTKREDEILNACKTOBERET 1026
QY 1023 KNLI-----FLYSGANIHOFSRFFEDKDAKESSKLLWIDKE-----D 1060
DQ 1027 ERLIKPILRGDKIKRYSWEWAGEWLINTHNGYTSNLK--SKIPPIDIEKYPATKAHLDAH 1084
QY 1061 LEKVUTK-----DNQOTERVFYRAIASNTNERTMTLSLPGNCYCNSIYIINDEKTPISL 1116
DQ 1085 YDTIATRCQDQGTPTHLRNCAY--LEDFFEKIVYTPETSQAGAYFIYENGSIKLEKTVFMI 1142
QY 1117 -----YKCLFIISFNSFVDFELLRFDVDSNVL-----KSLYOCMPQPEEKEILSN 1164
DQ 1143 VSDANLKLTLTALLSKLITFYFKNFCGCGCIIGKSGYQYNKHALEKIPQITERNQELA 1202
QY 1165 PLYLMLAKNTSLIARNDPENPKYLLYLEYFVKDEKVKNLKLDKEDFFKEKENENN 1224
DQ 1203 DKITDCAEQILAKADPKANTQKL-----EKEIDA 1233
QY 1225 IIASLYSLAKEDFITL 1240
DQ 1234 LVYQLYHLTDEBIKTI 1249

```

RESULT 4

```

US-09-738-626-4321
; Sequence 4321, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKAZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03

```


Db 1730 LDTE---KKONQMSIELDEKKONHORRIELDEKKONHMMHIELDEKKINKINILL 1786
 Qy 819 ---KNSIIAINNIYKFDKFTSGGDPNIFRYVFTNKLKIKKGNLTLYLVPASAINNESS 875
 Db 1787 BENKNKIIQINS-----CLOKEKEIN-----ELLNIOLEDKKNNIQEL---NIKLEDE 1831
 Qy 876 SRLRK-HIFARYKLANVIYO---FENKGRFKD---VHSFKAIFQLSNIK---ESTSS 924
 Db 1832 KKLNEQINTVLENKINNIQELNLIKLEDEKKLNEQININTLENKINNIQELNLIKLEDEKKON 1891
 Qy 925 FKAQMIQSSDNILKEITRDLKSDODAYKGIELNINOIKKLSPIQESIEIFKDNBEFTL 984
 Db 1892 QQINIQLBEEKNKIIQINSHLKKEKE---INEELNILENKKONIQEITIKLEBQKK--- 1945
 Qy 985 INKMFKFSALGEGYIDFKGLDPSIKNRKSLKKECNKVLIFLYSGANIHQNSRFPED 1044
 Db 1946 INE-----EQYIQLEK--DKEIIN--SMVMEKEKII-----NN---EI 1978
 Qy 1045 KDAKSSKLWIDKEDLEKVLTKDNQVQTERVVFYRAIASN--TNERWTISTLSPGNCYCV 1102
 Db 1979 KQKLEKEK-----KONDQVLIHLENEKQANKKLNILLDQNKINBEL----- 2020
 Qy 1103 NSIYINDEKTPISLYKKLFTIISFNFSVDFLLRRFVDSNLKSLCYQCPMPQPERKEIL 1162
 Db 2021 NIQVEQEK-----LINNEIIVQLAKENEENKINSLL-----EQON-- 2056
 Qy 1163 SNPLYNLAKNTSLIANDPENFKYLLYLE-----YKFDKEK---VNKILKLOK 1210
 Db 2057 -----GLNKKVTLQLEKEEENGKGLQLENEKQENGNLRPELENEKKDIANLILQLOE 2110
 Qy 1211 EDE-----FPKEKENENNFI 1226
 Db 2111 EKENTKNVMQMDKEKTKKNVMV 2134

RESULT 9

US-10-732-923-3340
 ; Sequence 3340, Application US/10732923
 ; Publication No. US20050108791A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Edgerton, Michael D
 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
 ; FILE REFERENCE: 38-15(52796)C
 ; CURRENT APPLICATION NUMBER: US/10/732,923
 ; CURRENT FILING DATE: 2003-12-10
 ; PRIOR APPLICATION NUMBER: 10/310,154
 ; PRIOR FILING DATE: 2002-12-04
 ; NUMBER OF SEQ ID NOS: 24149
 ; SEQ ID NO 3340
 ; LENGTH: 1939
 ; TYPE: PR1
 ; ORGANISM: Plasmodium chabaudi
 US-10-732-923-3340

Query Match 5.18; Score 335.5; DB 5; Length 1939;
 Best Local Similarity 20.18; Pred. No. 3.8e-10;
 Matches 304; Conservative 248; Mismatches 502; Indels 457; Gaps 75;
 Qy 18 KQSKDFIKKENINKLADF----FILJKNKLFSDNNSTANIESLLKIYFEELNYSVEQ 73
 Db 115 KQINKEYEKE--LNEKSEFIKQWELLKKEKELNLENKINKEIITLKREB----- 165
 Qy 74 QKAGQIEGVSVDILLFENDKDKASPNKKKEAKQNEPIPIBIDILIIAEVVRPTSPFD 133
 Db 166 ----KLNDIESE---VIEKNKEKELNYETNKMISLDKLTCB---VQEKK-----D 207
 Qy 134 AKDKLK-----ESEDQLYRINQYOKHY-----GILSGKWRLYDKSKVLYGKER-Y 180
 Db 208 NLEKINKVIEKENNL-RELKBFMEKQNEIIESLDGTINDKK--NAYEKLEISFEERKM 264
 Qy 181 IEFNFXKIEEBEYKEQEWFLTYLIRKERYLTSN--VIEVEKEQIAKEKEIIQKTLK 238
 Db 265 IEMLDKSLIEKEE-----NFANKQAKLSENEIIEKLKDIESKDFPSKEEK 313

Qy 239 EILYERPDSDIVFKIAKNIYDKE-FKLSGKEITTOHILASILEESIIFILRIFF----- 290
 Db 314 FASMENELNTLSKDSKACQMEVYKLEIKDLSQSL---VEKEREIPEIKNEYDDKINNM 370
 Qy 291 ----IAYIEDNDI-----FKILOENKLYRSSISFRYFFVDENTKKL----- 329
 Db 371 KEKLSINDKIGDNTVLHSEEBKINKLLKEKETELNEIHKYNLEIETIKNELNEKEBEL 430
 Qy 330 -EYKLIITI-----FNLLDKGSDAIFPVPNGSLFSEDKVKVYNNLEGLSISIEI 380
 Db 431 EKNKKAHTVEVNTLTKLEKKTEDAK---EG---HKNELNELNQ---LSKLNK- 478
 Qy 381 LVKMLFFPEKNIKDSEF-----VKYSRLDPKFSFGELYETL-----LEYDLRI 422
 Db 479 -----EKDNINKNTELNKDISLNSEVNLNKDQTLGNDIKTLNDLNNLKNKIN 531
 Qy 423 ADTTHRIIEGVIYLRTEEELENK--KVNKIATYLGKNIYLTSLRSLDRKSGAYTPPD 480
 Db 532 SDNKNKMKED--LAMLNEEMEGCVVIDEIEKKYKNEIFMLEEKLKEKENYA-----D 583
 Qy 481 LTDFMVISIEBQLTKSPLDIKIINDSCGSHFLISCLDYLTEKVKWYELDKFEDVK--- 537
 Db 584 LND-----EISILRNSI-----YVKEKEFIEMKEFYENKINL 615
 Qy 538 --KELDEEYRV-----IIEBSEYDVQDSISKELAVLKMLLKXCICYG 578
 Db 616 FKNFEEKKNYENELNSURLKYDNEQGLIKQIDELNIOKLKTEEYKLO-----LYND 668
 Qy 579 DINPISVEITMLSLMINTPIFGTPL-SFIEHHK-----AGNALLGYTKDEPDIYKVKKF 632
 Db 669 NMHFRSICTKIDMPYSENIKGSDLVDVFTAVIKRDESSDANPDPTTHKEMVAELEKRH 728
 Qy 633 ESGFSLFKRKRIKEITIL-----EDIYQIKGINDTTKEDIEKSKKIY 675
 Db 729 AAIVAELEKHEKEIAKLGEHKEVVLRLGEQHEKEETIILEEKHKDQVVTKLGEQHKENI 788
 Qy 676 KEYEESKDIDLRIIFSLIKLYSLSPDKSLNMFSDIASVISIENILNKTSSDEKEI 735
 Db 789 KLEEEHKOV-----VTKLDQYKEEIAKUEEHAVVVAELEK---HKLGEHKEW 837
 Qy 736 EKIRK-----LSYYKFFHYGIEFPDIOBGFDIVIGNPWEKTKFNETE----- 779
 Db 838 DELEKHEADPVEGLEBKH-----AETAKLBEG-----HKSEMNEVEKRHADFVE 882
 Qy 780 -PFSKHIPNRYKLGIKENI-----KQEI--LSKNHPLSIEYNEE 818
 Db 883 GLEEKHKAETAKLGEHREVAVAGLEEKHEVVAELEKHEEIAKLEEGHEKYMALGEK 942
 Qy 819 KNSIIAINNIYKFDKFTSGGDPNIFRYVFTNKLKIKKGNLTLYLVPASAINNESSRI 878
 Db 943 HKEVWA-----GLEAKHNLBEGHEW-----ALEKRHADLV-----AVLEEQHKAEL 986
 Qy 879 LR-----KHIFA-----RYKLNVIYQENKK-----RPFKVH----- 905
 Db 987 IKLGEHEKEVWAGIEBEKYVEALKAEHKKDQVVTKLGEQHEKEIAKLGDGHEKVVNEVEK 1046
 Qy 906 -----SSPKFAIFOL-SNIKESTSFKAQFM-----IOSSNLIKKEITRDLKDSK 949
 Db 1047 KNASLLNMLEENHNKEMIKLEEHKESASDLVEKLYQKOBEEVKNKINBELTNVINDLN 1106
 Qy 950 D-----DAYKGIELNINOIKLSPIOBSIIEFKDNEEFTLINKMFSFALSALGEG 998
 Db 1107 DSMCYKQIILEVEVEKREYN-EEINKLKIVQN---EMKDNNDKKILEK-----EN 1153
 Qy 999 YIDFKGLDPSIKNRKSL-LKECNKNLIFLYSGANIHFNSRFFEDKDAKESKLWID 1057
 Db 1154 EI---KLNKLSNYKVFETKENTYKN-----SEMVVNENKRIIVDSVCKEN-----IS 1200
 Qy 1058 KEDLE-----KVLTKDNQVQTERVVFYRAIASNTNERT-MISTLSPGNCYCVNIYINDEK 1111
 Db 1201 ESDVEGKGGLNKMTLSLKKERNIF--SINDKNSESSELVDT-----IKSAYIN--- 1247

QY 1112 TPISLYKLF-----IISFNSFV-FDELLRRFVDSNVLSKSLYQCPQPE 1157
DB 1248 -KIEMTKEDNGKNIEDLNKLDLSNBLINLNKNVLTDDNNLNKKEBIKONKN 1306
QY 1158 EKEILSNPLYNLAKNTSLIAKND-----PENFKYLLYLEVPKPKKEKVKNI 1205
DB 1307 EKEKNTEILNL--NDDIIFLKEISEWKDEBEKLTENIK-----LKNDEIQINKE 1357
QY 1206 LKLDKDEPFYKVENNNFIASL---YSLAKEDFTLLGDFAKLNKKKGEDYISL 1262
DB 1358 YKIEENMLTKFENINE--VTSLNQIEIKMKLBELKNVYLLAEKR-ETNMS--IS 1412
QY 1263 GYDNYLLNKKI 1273
DB 1413 NDNKLVENNI 1423

RESULT 10
US-10-732-923-3354
; Sequence 3354, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10732,923
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3354
; LENGTH: 1527
; TYPE: PRP
; ORGANISM: Plasmodium yoelii yoelii
; NAME/KEY: unsure
; FEATURE:
; LOCATION: (1)..(1527)
; OTHER INFORMATION: unsure at all xaa locations
US-10-732-923-3354

Query Match 5.0%; Score 328.5; DB 5; Length 1527;
Best Local Similarity 21.3%; Pred. No. 7.1e-10;
Matches 247; Conservative 210; Mismatches 386; Indels 319; Gaps 61;

QY 1 MKNIDIVKTNPNISLYKQSKDFIKKENINKLDFILIKNKLFSIDNSTEANIESLL 60
DB 544 MDENPIILDSNTIN-----KNIIIDAESDKLVQNLDMIKIL--QDDKN--NIDTII 594

QY 61 KYIFBELNYSVEQKAGQI-----EGVESRVDILLFEN-----DKOKASFN 101
DB 595 KNESEKIT-SLLEKTSMEILCNKICQLKKKQVEDKNDTIREETHQINIDKKEBI 653

QY 102 NKLKAKONNEPIEDILLIAEVKPTFSFDAKDKLSESDQLYVLYNQYKHGILSN 161
DB 654 NNLLYV-----INIKNKIKSELEAFTNSFNLKLLKMKNDBI--INIKRNVNFLT 704

QY 162 GKWRLYDKSKVLYGKRYEFNFXXKIEEKEBYKEQEWFLVLYIRKERYLKTNSV-- 218
DB 705 -TINKLEEKRL-----KYAEFK-----EKEKYKE-----LIMKKEELLIXQINNK 746

QY 219 --TEVEKEQIAKBEIQQ--TLKEILYERPDSDIVFKIAKNYDKFKLSGKEITQHI 273
DB 747 KNIEFTKQOQMLKDEIQIETLKTKEIEBLKQD-----FEKTQKINMENLEME- 795

QY 274 LASILEESIIFILRIEIAVIBNDI--FKKILQ--ENKLYRSSISFRYFFVD--ENTK 326
DB 796 -----RES-----FINNNLENEINMKKELEEKYETQIKETMKYQIKEEIEKTK 842

QY 327 KLEYKXIIITIFNLLDKGSDAIFKPPVNGLSEDKVKYLNNEGLLSISBIEI----- 380
DB 843 QNAEQN-----FN-----SRPEKYKEML--EKNKNDFINNLLIEKNNEIESFKNDIEQ 888

QY 381 --LVKMLFFBEKNIK-----DEKFKVYSRLDPKPSFGELYETL-----LEY----- 418

DB 889 KKFEMEKKEBENKLLQNNFNEMKNHFIEEQNKHIENIKKEYELIKNEIEYLYKEEMKK 948
QY 419 -----DLRIADTTVHRIIEDGVYLIRTEBEE--NKKVKIATYKGNIVYITSRS 466
DB 949 NKIQEIEVENELKAD-----EKNKHIDMKKELEINVEINNLKEKINAEQTQAQEM 1001
QY 467 LORUKSGAYTTPDDLTDFWVI--SSIEEQK-TKSPLDIKIID--NSCSGHFLISCLD 520
DB 1002 LETQR-----NEVEQKLEIQORNEYEQOLEIRRNEMENKLEIQRKESYDLGSKKESEIQ 1054
QY 521 YLTKVWYELDFEDVKKELDEYRVIIIESSEYDVODSIS-----KELVLK- 568
DB 1055 ILT-----DKIEKLEIE--IINKKIDLEHISILNDRERILNQNKULEDN 1101
QY 569 MLLKXCIYGVINPISVEITMLSLWNTFIFGTPLSFIEHHIKAGNALLGYTKDEFFDIV 628
DB 1102 LKNNEIYNKINILNDEKLE-----KEIEHKEGAKESDQIREQPADLL 1149
QY 629 KKKFSGFLFKRIKEIITIL-----EDIYKIKGINDTTK-----EDIEK 670
DB 1150 QBEIDR-----IKKESKEKVIYIQVYNEINEEYEEKKEYNLDLEKAVNSKQLTKECEE 1205
QY 671 SKKIYKEYES-KDIDNLRILFSLIKLYSL-----SFKSLNMFSDIASVISLIE 720
DB 1206 NIQKINEYEDMIKMLENQTENVLTKIQBELNEDFLKKEAFNERNKNDLLKNYEHAIENK 777
QY 721 NI-----LGKNTSSEDEKIRKIRKLSYYKFFHYGIFPDIOGFDIVIGNPPWEKTKFNE 777
DB 1266 HIKEQLENFTNS-NEEKISQIKN-----QYETQIKEMQIHFDL-----NETKNKN 1311
QY 778 TEFFSKHI--PNYRKIGIKQNIIOKILSKONHPLSIEYNEEKNSIIANNIYKDFPKC 835
DB 1312 NEMLEDIVKEQNDKEQSIQHMNSVNTLVTK-NEQLNLNITIKNEL--SNFOEKYD--- 1365
QY 836 FTSGGDPNLPFRYFVFNKLKIKKGNLTYLVPSAINWNESSRILRKHIFARYKLVYQF 895
DB 1366 -----KLVKENHLK-----NEBQTFPLDKILDLENVNRKLMLEV 1399

QY 896 ENKCR---FKDVHSGFKPAIFQLSN--IKESTSFKAKFMIOQSDNLIKETIRLDKSKD- 950
DB 1400 TEKEREINKNV-----IKSLQIQINEQTKLYK-EYTDLKDDEL-QLKKQKNSIDS 1452
QY 951 ----DAYGIELNINQIKLSPIQBSIEFKONEEFTLINKMFKSPFALGEGVIDPKKGL 1006
DB 1453 NVPIDANSSINNALLEENKKLQNEVISTKNDTMTVDNLTKRLS-----PIESK--- 1504

QY 1007 DSIKNRK--SLLKECNKNL 1025
DB 1505 ---IKEEKYGNDIIOQANDLQ 1523

RESULT 11
US-10-732-923-3351
; Sequence 3351, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10732,923
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3351
; LENGTH: 1985
; TYPE: PRP
; ORGANISM: Plasmodium yoelii yoelii
US-10-732-923-3351

Query Match 5.0%; Score 324; DB 5; Length 1985;
Best Local Similarity 19.7%; Pred. No. 1.8e-09;


```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...740
; SEQUENCE DESCRIPTION: SEQ ID NO: 7398:
US-10-335-977-7398

Query Match          4.9%; Score 318; DB 4; Length 740;
Best Local Similarity 24.1%; Pred. No. 1.1e-09;
Matches 176; Conservative 99; Mismatches 240; Indels 214; Gaps 35;

QY 22 KDFIKEN-----INKLKDFILIKNKLFSIDNSTEANIESLLK--YIFEELNYS 70
DB 54 KVFIKYNQEPKPKETIENFEKEINSLEENVPRQDDEEQKNEIFLNKAYGDCNTY- 112
QY 71 VEQQKAGQIEGVSVDILLFPENDKDKASFNKLKAKQKNNEP-----IEDI- 119
DB 113 ---EKVDSVIYDKEVRVLI-----EVKALNNKTFPKNRENPLSKAFQCMVLYFLEET 164
QY 120 -----LIIAEVVRPFSFDKDKLKESED-QLYRYLQYQKHVGLSNGKVRWLYD 169
DB 165 KKNLSKHTIICNA-HEFFLPCKDOLLSDKEDRIKKFYKNYAKKEGT-----D 213
QY 170 KSKVLY-----GKRYIEFNFXKIBEEKEEYEQEWFVLYFYLIRKERYLKTS 216
DB 214 SSKPKFYKLEQPLQDFQGLRYTHFNL-----SDDFKE--LPLIYQVLSQVLLKQK 265
QY 217 NVIEVEQIAKEKEIIQTKLEILYERDDD--SIVFKIA--KNYIDKGFKSGKEITH 272
DB 266 RTLDAN-----TLNKDFYEELLYILGLEQNDKGLIKPSRTKNSLSGALKERYKNLDB 321
QY 273 ILASILEESIIFILRIFFATYIEDNDIIFKKILOENKLYRSSISFRYF-----FYDENYK 326
DB 322 EYNALL--IANNRILFLRLLESLLISFEHFEKERSFUTGNFKDFNALNTLFFEVLA 378
QY 327 KLEYYKKIITIFNLDKGSDAIFPVPFNGGLFSE-----DKVKYLNNEGLLSISEIEE 379
DB 379 KNSERLPEIKENKILG-----KIPYNSLSLFDKTPLELKGHEIKLLDNKKL-----E 425
QY 380 ILVKMLFFEEKNIKDKFKVYSRLDPKSPGELYETILEY-----DLRIAD-TTVHRIEDGV 435
DB 426 IYKNSVLKCHENYQDK-----ASLPLLEYFFKFLRVYDFTTTPKIDKWT 471
QY 436 YLIRTEEELENKWKNIATYLGNIYLTSLRDKKSGAYYTPDDLTDFWVSIIE-----491
DB 472 --DTSESRLINPSV-----LGLVFEKNGYKEGFSYTPSPITSYNCKESITTVIL 519
QY 492 -----EOLKT-----KSPLDIKIIDNSCGSHFLISCLDYL 522
DB 520 DFNQKNVTECDKLEKKNYKNSYKEDKKRYLQLLLTIRICDPVAGSGHFLVSVLN-- 577
QY 523 TEKWV--YELDKPE-----DVKKELDEYRVIIIEES-----BEYDQVDSISKE 563
DB 578 -EMVWAYKGLIASLYRYDLKLENDE--IIHYPTGTGEI FNYTIPHSENDPHHIOXE 633
QY 564 LV-LKRMMLKCKYGVNDINPISVEITMLSLWI-----NTFPGTPLSFIEH 608
DB 634 LPELKKSIIENCLFGVDINPNSCEITKLRLWIELLYKSYVYIFBEGKNTNALET-LPNIDI 692
QY 609 HIKAGNALL 617
DB 693 NIKCANSLI 701
```

RESULT 13

US-10-954-924-7

; Sequence 7, Application US/10954924

; Publication No. US2005095256A1

; GENERAL INFORMATION:

```
; APPLICANT: Bujard, Hermann
; APPLICANT: TOLLE, RALF
; APPLICANT: PAN, WEIQING
; TITLE OF INVENTION: Recombinant Process for Preparing a Complete Malaria Antigen GPI
; FILE REFERENCE: GRUB-003DIV
; CURRENT APPLICATION NUMBER: US/10/954,924
; CURRENT FILING DATE: 2004-09-29
; PRIOR APPLICATION NUMBER: 09/269,874
; PRIOR FILING DATE: 1999-08-02
; PRIOR APPLICATION NUMBER: PCT/EP97/05441
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: DE 19640817.2
; PRIOR FILING DATE: 1996-10-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1602
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-954-924-7
```

```
Query Match          4.8%; Score 316; DB 5; Length 1602;
Best Local Similarity 19.7%; Pred. No. 3.9e-09;
Matches 295; Conservative 229; Mismatches 481; Indels 490; Gaps 73;

QY 2 KTDI-----VKTNNPNISLYKQSKDFIK-----KENTINKLKDFILIKNKL 45
DB 172 KLNDVCANDYCOIPFNLKITRAMELDVCLKVGYRKPDLNIIKDNVGMEDY--IKNK-- 227
QY 46 SIDDNSTEANSLSLKYIFEELNYSVEQQAQIEGVESRVDILLPENDKDKASFNKLK 105
DB 228 -----KTIENTINELI-----EESKKTIDKNKNAATKEEKKK-----LYQAQYDLSIYNKQLE 274
QY 106 EA-----KKNNEPIEDIL-IIAEVKKR-----PTSFDAKDKLKE 141
DB 275 EAHNLISVLEKEDIDLKKNEN--IKELDKINEIKNPPANSNGTPTNLLDKNKKIEH 331
QY 142 E-----DQLYR-----YLNQYQKHVGLS-----NKKVVR 166
DB 332 EKXEIKIAKTIKFNIDSLFTDPLELEYLYRKNKNIDISAKVETKESTEPNPGVTYP 391
QY 167 L-YDSKVLVYGEKRYI-----EFNFKIEEKEEYEQEWFVLYFYLIRKERYLKTSNVIE 220
DB 392 LSYNDINNALNELNSFGDLINFPDYTKBPSKNIYTDNE-----RKKFINEIKKIK 442
QY 221 VEKBOIAKEKEIIQ--KTLKEILYERDDSVIFKIAKNYDKEF-----KLSGK 267
DB 443 IEKKKIESDKSYEDRSKSLNDITKEYE-----KLNIEIYDSKFNNNIDLTNFEKMGK 496
QY 268 -----EITQHILASILEESIIFILRIF-PIAYIED-----NDIFPKILOENKLYRSSIS 315
DB 497 RYSYKVKELTHNTFASVENSXHNLEKLTALKWMEDYSLRNIIVBEKLYYKNLISKIE 556
QY 316 FRYFPVDENTKK--KLEYKKIITIFNLDDKG-----SDAIKPPVFNGLFSEDKVKYLVN 368
DB 557 NEIETLVENIKDEEQLEPKKITDKNKPKDEKILEVSDIVKQV-----600
QY 369 EGLISISEETELVKMLPFEKNIIDKFKVYSRLDPKSPGELYETLLEYD--LRIADT 425
DB 601 QKVLLNKKIDELKTKQILKONVELKHNHVPNSYQENKQEPYLYLIVKKEIDKLKVFMP 660
QY 426 TVHRIEDGVYLIRTEEELENKKNVKNKIATYLGNIYLTSLRDKKSGAYYTPDDLTDFM 485
DB 661 KVESLINEEKNIKTEGQSDNSEPS-----TEGEITGQATTKPGQAGSALEGDS-----710
QY 486 VISSIEEQLKTSPLDIKI-----IDNSCGSHFLISCLDYLTEKW-----527
DB 711 VOQAQEQKQAQPPVPVPVPEAKAQVPTPPAPVNNKTEN-----VSKLDYL-EKLYEFLNT 765
QY 528 -----YELDKFEDVKKELDEYRVIIIESEYVDQDSISKELVLKMLKXKCIYGVN 591
DB 766 SYICHKIIVSHSTWNEKILKQYKITKEESKLSLSCDPLD-----LLFNQI-N 812
```

```
QY 582 PISVEITMLSLWINTFPGTPLS--FIEHHIKAGNALLGYTKDEFFDIVKMKFESGFSLF 639
Db 813 NIPVMYSMPDSLNS-----LSQLFWEIYKEWVCNLYKLDN--DKIKNLLEA---- 860
QY 640 KRIKEIITILEDIYQIK-----GINDTKE-----DIEKSKKIY 675
Db 861 KKVSTSVKTLSSSSMQPLSLTPQDKPEVSANDDTSHSTNLNLSKLFLFENILSGKNKIY 920
QY 676 KEYESKIDNL--RIIFSILKLYSLSPKSLNMFSDIASVISLLENILGNKTSSEDK- 732
Db 921 QELIGQKSENFEYKILKDSDTFYNESFTNFVSKADDINSL-----NDESKRKL 971
QY 733 -EKIEKIRKLSYYKFFHYGIEPDIQEGFDIVIGNPPWEKTKFNETEFPSKH--IPNYR 789
Db 972 EEDINKLK-----TLQSPDL-----YNYKKLKLRLFDKKTGVKY- 1009
QY 790 KLGIEQNIIOBILSKDNHPLSIEYNEEKNSIIANNIYKDFPKFTSGGDNLPFRYFV 849
Db 1010 KMQIKKLTLLKEQLESKLN-----SLNN----- 1032
QY 850 TPNLKLKEGNTLYVPSAINWESSRILRKHI PARYKLYNIYQFENKKRFDVHSSFK 909
Db 1033 -----PKHVQNFSV-----PFNKKEAEIAET-- 1055
QY 910 FAIPQLSNIKESTSSFKAKFMIOSSDNILKEITRDLKDSKDAYKGIELNIN-OIKKLSP 968
Db 1056 -----ENT-----LENTKILLKH-----YGLVKKYNGESSPLKT 1085
QY 969 IOBSIIIEFKDN-----BEFTLINMKFSALSGEYIDPK-----GLDPSIKNRKSL 1016
Db 1086 LSEESIQTEDNYASLENFVLSKLEGL-----KDNLNLEKKLSYLSGLHLHIAELKEV 1141
QY 1017 LKCNKNKLLFLYSGANIHFOPNSRFPEDDKAKESSKLLWIDKEDLEKVLTKD-----NQY 1071
Db 1142 IK--NKN-----YTGNSPSENNT---DVNNALSEYKFKPLPEGTDVATVWSESQDTLEQS 1191
QY 1072 QTERVFRATASNERTMTISTLSPGNCYCNSIYINDERTPTISLYKGLFIIGFNSFV 1131
Db 1192 QPKKPASTHVGAESNTIT-----TSQNVDEDDV-----IIVIFGESEE 1232
QY 1132 DF-----LLRREVDNSVLK-----SCLYQCPMP---QPEKEILSNPLYNLAK 1172
Db 1233 DYDDLQGVVTGEAVTPSVIDNLSKIENEVEVLYLKLPLAGVYRSLKQLENNWMTFNVV 1292
QY 1173 NTSLLIAKDNPENPKYLLYLEYKFP-DKEKNKILKLDKEDEP-FKEKENNNFIASLY 1230
Db 1293 KDILNSRFNKRNFKNVLESDLIPYKDLTSSNYVVK---DPYKFLNKRKDRKPL--SSY 1346
QY 1231 SLAKEDFIT-----LLGDFKALKNKKGE-----DYISSLIKGYDNYL--LNN 1271
Db 1347 NYIKOSITDINFANDVIGYKILSEYKYSKSDLSIKKIYINDKQGENEKYLPFLNN 1401
```

RESULT 14

```
US-10-954-924-5
; Sequence 5, Application US/10954924
; Publication No. US20050095256A1
; GENERAL INFORMATION:
; APPLICANT: Butjard, Hermann
; APPLICANT: TOLLE, RALF
; APPLICANT: PAN, WEIQING
; TITLE OF INVENTION: Recombinant Process for Preparing a Complete Malaria Antigen GPI9
; FILE REFERENCE: GRUB-003DIV
; CURRENT APPLICATION NUMBER: US/10/954,924
; CURRENT FILING DATE: 2004-09-29
; PRIOR APPLICATION NUMBER: 09/269,874
; PRIOR FILING DATE: 1999-08-02
; PRIOR APPLICATION NUMBER: PCT/EP97/05441
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: DE 19640817.2
; PRIOR FILING DATE: 1996-10-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 5
; LENGTH: 1621
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-954-924-5

Query Match 4.8%; Score 316; DB 5; Length 1621;
Best Local Similarity 19.7%; Pred. No. 3.9e-09;
Matches 295; Conservative 229; Mismatches 481; Indels 490; Gaps 73;

QY 2 KTNDI-----VKTNNPNISLYKQSLKDFIK-----KENINKLAKDFILIKNKL 45
Db 191 KLVNDVANDYCOIPFNLRANELDLVKLVFGYRKPLDNIKDNVGMEDY--IKNK-- 246
QY 46 SIDDNSTEANIESTLKYIFEELNYSVBQKQACQIGVESRVILLFENDKDKASPNNKLK 105
Db 247 -----KTENINELI-----EESKKTIDKKNATKEEEKK-----LQAOYDLSIYNKQLE 293
QY 106 EA-----KKNNEPIEDIL--IIAEVKK-----PTFSFDAKDKLKEBS 141
Db 294 EAHNLISVLEKEIDTLKKNEN--IKELLDKLINEIKNPPPPANSNGTPTNLLDKNKKIEH 350
QY 142 E-----DQLYR-----YLNQYQKHVGLS-----NGKVR 166
Db 351 EKEIKEIAKTIFNIDSLFTDPLEYLYLRKKNKIDISAKVETKESTBPNEYPNGVTYP 410
QY 167 L-YDKSKVLYGKRYI-----EFNFKIIBEKEEYKEQWFLVFIYLIRKERYLKTNSVIE 220
Db 411 LSYNOINNALNELNSFGDLINFPDYTKEPKNIYTDNE-----RKFFINEIKEKIK 461
QY 221 VEKEQIAKEKEIIQ---KTLKEILYERPDSDIVFKIAKNIYDKEF-----KLSGK 267
Db 462 IEKKIESDKSYEDRSKSLNDITKEYE-----KLLNEIYDSKFNNDILTNFEXWGMGK 515
QY 268 -----EITQHILASILEESIIFILRIF-FTAYIED-----NDIFKIILOENKLYRSSIS 315
Db 516 RYSYKVKELTHNTFASYSKHNLEKLTALKATYMEDYSLRNIVVEKELKYKNLISKIE 575
QY 316 FRYPPFDENTKK---KLEYKKIITIFNLDDKG---SDAIKFPVFNGLFSEDKVKYLYNN 368
Db 576 NIETLVENIKDEEQLEKFKITKENKPEKILEVSDIVKQV----- 619
QY 369 EGLISISEIEEILVKMLFPEEKNIKDEKVKYSRLDPKSGELIYETLLEYD---LRIADT 425
Db 620 QKVLLANKIDELKKTQILKQVLEKHNIHVPNSYKQENKQEPYLLIVLKEIDKLKVFMP 679
QY 426 TVHRIIEDGVYLIRTEBELENKVNKIATYLGNIYVLTSRSLDRKSKSGAYTTPDDTLDFM 485
Db 680 KVESLINEEKNIKTGEGQSDNSEPS-----TEGEITGOATTKPGQOAGSALEGDS----- 729
QY 486 VISSIEQLKTKSPLDIKI-----IDNSCGSGHFLISCLDYLTEKVV----- 527
Db 730 VQAQOEQKQAOFPVFPVPEAKAQVTPPPAPVNNKTEN-----VSKLDYL-EKLYEFLNT 784
QY 528 -----YELDKFEDVKKELDEEYRVIEESEYDVQDSISKELVLKMLLKXCIYGV DIN 581
Db 785 SYICHKYILVSHSTWMEKILQYKITKEESKLSKSCDPLD-----LLFNIQ-N 831
QY 582 PISVEITMLSLWINTFPGTPLS--FIEHHIKAGNALLGYTKDEFFDIVKMKFESGFSLF 639
Db 832 NIPVMYSMPDSLNS-----LSQLFWEIYKEWVCNLYKLDN--DKIKNLLEA---- 879
QY 640 KRIKEIITILEDIYQIK-----GINDTKE-----DIEKSKKIY 675
Db 880 KKVSTSVKTLSSSSMQPLSLTPQDKPEVSANDDTSHSTNLNLSKLFLFENILSGKNKIY 939
QY 676 KEYESKIDNL--RIIFSILKLYSLSPKSLNMFSDIASVISLLENILGNKTSSEDK- 732
Db 940 QELIGQKSENFEYKILKDSDTFYNESFTNFVSKADDINSL-----NDESKRKL 990
QY 733 -EKIEKIRKLSYYKFFHYGIEPDIQEGFDIVIGNPPWEKTKFNETEFPSKH--IPNYR 789
Db 991 EEDINKLK-----TLQSPDL-----YNYKKLKLRLFDKKTGVKY- 1028
```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:59:25 ; Search time 20.4381 Seconds
(without alignments)
633.203 Million cell updates/sec

Title: US-10-688-058-16
Perfect score: 6538
Sequence: 1 MKTNDIVKTNPNISLYQL.....SSLIQYDNYLLNNKIFTHK 1277

Scoring table: BLOSUM62

Searched: 70609 seqs, 10134256 residues

Total number of hits satisfying chosen parameters: 70609

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:
1: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep:
2: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep:
3: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep:
4: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep:
5: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep:
6: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep:
7: /cgn2_6/prodata/1/pubpaa/US11_NEW_PUB.pep:
8: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	284.5	4.4	2665	7 US-11-124-368A-214	Sequence 214, App
2	269	4.1	2710	7 US-11-051-453-41	Sequence 41, Appl
3	266	4.1	2668	7 US-11-124-368A-215	Sequence 215, App
4	261.5	4.0	1786	7 US-11-196-400-3	Sequence 3, Appli
5	224	3.4	1652	6 US-10-995-561-663	Sequence 663, App
6	224	3.4	1938	6 US-10-995-561-662	Sequence 661, App
7	224	3.4	1169	7 US-11-077-550-20	Sequence 20, Appl
8	222	3.4	1938	6 US-10-995-561-662	Sequence 662, App
9	219.5	3.4	1954	6 US-10-995-561-660	Sequence 660, App
10	219.5	3.4	1972	6 US-10-995-561-664	Sequence 664, App
11	219.5	3.4	1972	6 US-10-995-561-666	Sequence 666, App
12	219.5	3.4	5024	6 US-10-793-626-2594	Sequence 2594, App
13	211.5	3.2	1404	6 US-10-878-556A-169	Sequence 169, App
14	206.5	3.2	885	6 US-10-793-626-1660	Sequence 1660, App
15	206.5	3.2	1070	7 US-11/062	Sequence 4, Appli
16	206.5	3.2	1095	7 US-11/062	Sequence 7, Appli
17	205.5	3.1	2367	7 US-11-051-453-42	Sequence 42, Appl
18	198.5	3.0	1614	6 US-10-821-234-903	Sequence 903, App
19	192	2.9	1147	7 US-10-615-668-5	Sequence 5, Appli
20	190.5	2.9	976	7 US-11-155-288-20	Sequence 20, Appl
21	188	2.9	1960	6 US-11-069-834-48	Sequence 48, Appl
22	187.5	2.9	1216	6 US-10-873-528-12	Sequence 12, Appl
23	182.5	2.8	1976	7 US-11-069-834-54	Sequence 54, Appl
24	180	2.8	1315	7 US-11-077-550-141	Sequence 141, App
25	178.5	2.7	708	7 US-11-196-475-76	Sequence 76, Appl

ALIGNMENTS

RESULT 1

US-11-124-368A-214
; Sequence 214, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 214
; LENGTH: 2665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-368A-214

Query Match 4.4%; Score 284.5; DB 7; Length 2665;
Best Local Similarity 20.0%; Pred. No. 2.6e-07;
Matches 291; Conservative 224; Mismatches 454; Indels 489; Gaps 73;
QY 1 MKTNDIVKTNPNISLYQLSKDFIKENINKL-----KD 35
1266 INTDLEKSHTKQBEIPVLHBEQELLPNV---KEVSE---TQETMNELELLETSQSTTKD 1319
36 FILLINKLFSIDNN-----STEANIESL-----LKYIPEELNYSVEQKAGIEGVESR 85
1320 STTLARIEMERLRLNEKFSQSEIKSLTKERDNLTKTKEALEVHKDQK-----EHRET 1375
86 VDLIFENDKKKASFNKKLKEAKKNEPIEDILIAVKKPTFSFAKD----- 136
1376 LAKIQESQKQBSLNMK-----EKDNETK-----IVSEMEQ-----FKPKDSALLRIET 1422
137 -----KLKESDQLYRYLNQKHGYLSNGKWRLYDKSKVLYCEKRYIBFNFKKIE 190
1423 MGLSKRLQESHDEMKSAKEDD-----LQRLQEVLOSQDLKENIKEIVA 1470
191 KEEYKEQEWFLFYILIRKERYLTKSNVIEVEKQIAKEKEIIQKTLKEILYERPDSDIV 250
1471 KHLTEEEELKVAHCCLKEQE-----ETINELRVNLSEKETEISTIQKLEAI-----NDKIQ 1522

QY 251 FKIANIYDKPKLSGKBITQHILASILEESIIFILRIFIAIYEDNDIFKKILOENKLY 310
DB 1523 NKI-QBIYKBEQFNKQISE-----VOEKV-----NEL-KQFKHEHRKAK 1560
QY 311 RGSISPRYFYDENTKKKLYKIIIFNLLDKGSDAIFKPPVFNGLFSEDKVKVLLNNEG 370
DB 1561 DSAL-----QSTI-SKMLELTRLQESQEIQIMI-----KEKEMKRVQE 1600
QY 371 LLSI-----SEIEILVRLPFEEKNIKDEKPVKSRDLPKPSFGBLYETLLEYDLRIAD 424
DB 1601 ALQIERDLQKENTKEIVAKM--KESQEKYQFLKMTAVN-----ETQKMCIEI----- 1647
QY 425 TVVHRIEDGVILRTEBLENKVNKIATYLGNIYLT-----RSLDRKSGAY 475
DB 1648 -----HUKQEFQKLN--LENIETENIRLTQILHENLEEMRSVTKER----- 1688
QY 476 YTPDDLTDFWISSIEQKTPSLDIKIDNSCGSHFLISCLDYLTETKVVYELDKPBD 535
DB 1689 ---DDL-----RSVETLKVERD---QLKEN-----LRETIITRDLKQEE 1722
QY 536 VK-----KELDEYRVIBE-----SEYDVOD--SISKELVLKRLM 570
DB 1723 LKIVHMLKHEQHTDKLRGIVSEKTEISNMQKLEHNSDALKAQDLKIQEBELRAHMH 1782
QY 571 LKXCIVGVD--INPISVEITMLSWINTFIFGTPLSFIEHIIKAGNALLGYTKDFPDIVK 629
DB 1783 LKEQETIDKLRGIVSEKT-----DKLSNMQKLENSAKLQEKIQL----- 1825
QY 630 KKFESGSLPKKRIKEI---ITILEDYOKIKINDTTK-----EDIEKSKTYKYEBEK 682
DB 1826 KANEHQLITLKUVNETQKVSEBQKQIKQOSITLSKLEIENLNLAKQLEHLEEMK 1885
QY 683 DI-----DNLRIIFSLIKL-----YSLSPDKLSNMEFSDIASVISIENILGNKTSSEBK 732
DB 1886 SVMKERNLRRVEETLKLBRDLKESLQETKARDLEIQELKTARML-----SKEHK 1937
QY 733 EKIEKIRKLSYKPHYHGYIEFPDIQEGPD-----IVIGNPPWE 771
DB 1938 ETVDKLRE-----KISEKTIQISDIQKDLKSKDELQKTIQELQKELQLLRVKEDVMS 1992
QY 772 KTFNTEPFPSKHPNVRKLGIGKEONIKQELLSKNHPLSIYNE-----EKNS 821
DB 1993 HKKINEMEQKKQ-----FEAQNLSMQSV--RMDNFQTKKLHESLEIRIVAKERDE 2043
QY 822 IIAINNIYKPDFKFTSGGPDNLFYFVTFNLKLIKKEGN-----LTYLVPS 868
DB 2044 LRRIKESLKWED-----QFTATLREMIARDQNHQVQKPEKRLSDCQOHLTES 2092
QY 869 AIWNESSRILRKHI FARYK-----LNYI-YQPENKKRFPKDVHSSKFPAIFOLSNIX 919
DB 2093 L--REKCSRI--KELLKRYSEMDDHYECLNRLSLDLEKETEFHRIMKGLKYVLSYVTKIK 2148
QY 920 -----ESTSPKAFM--IQSSDNLKBEI-----TRDKOSKODAVKGIELMINQI 963
DB 2149 EBOHESINKEMDFIDEVEKQKELLIKIQLHQDCDVPSELKD-----LKLQN 2198
QY 964 KKLSPQESIEIFKONBEFTLINNMFSKFGALGEGYIDFKGLDPSIKNRKSLKBECKNK 1023
DB 2199 MDLH-IEEILKDFSES--EPPSIKTEFOQV-----LSNRKEM----- 2232
QY 1024 NLIFLYSGANIHPNSRFFEDKAKESKLLWI-----DKEDLEKVLTKQNG--YQTERVF 1077
DB 2233 -----TOFLEE-----WLNTRFDIEKLNGIQKENDRICQVNNFF 2267
QY 1078 YRAIANTNERTMTSLSPGNCYCVNSIYINDEKTPISLVKCLFIISFNSFVDFDLRR 1137
DB 2268 NNRIIAMNISTEPEERSA-----TISKEWQDLKSLKEKNEKL----- 2306
QY 1138 FVDSNVKLSCLYQCPMPQPEKEITLSNPLYNLAKNTSLIAK--NDPENFKYLLYLEYFK 1196
DB 2307 FKNYQTLKTSLASGAQVNPFTQD--NKNPHVTSRA--TQLTATEKIRELENSLH-----EAKE 2359

QY 1197 FDKEKVNKILDKDEBFFKEKENENNFIILASLYSLAKE-----DFITLLGDFKAL 1247
DB 2360 SAMHKESKLIIRKQKELEV-----TNDIIAKQAKVHESNKCLEKTETIQVLQDKVAL 2412
QY 1248 ---KNVKKGEDYVSSLIK 1262
DB 2413 GAKPYKEBIEIDLKMKLVK 2430
RESULT 2
US-11-051-453-41
; Sequence 41, Application US/11051453
; Publication No. US20050287150A1
; GENERAL INFORMATION:
; APPLICANT: AMEROSINO, DONNA
; APPLICANT: BABCOCK, GREGORY J.
; APPLICANT: BROERING, THERESA
; APPLICANT: GRAZIANO, ROBERT
; APPLICANT: HERNANDEZ, HECTOR JAVIER
; APPLICANT: LOWY, ISRAEL
; APPLICANT: MANDELL, ROBERT
; APPLICANT: MOLRINE, DEBORAH
; APPLICANT: THOMAS, JR., WILLIAM D.
; APPLICANT: ZHANG, HUI-FEN
; TITLE OF INVENTION: ANTIBODIES AGAINST CLOSTRIDIUM DIFFICILE TOXINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: MJ1-001
; CURRENT APPLICATION NUMBER: US/11/051,453
; CURRENT FILING DATE: 2005-02-04
; PRIOR APPLICATION NUMBER: 60/542,357
; PRIOR FILING DATE: 2004-02-06
; PRIOR APPLICATION NUMBER: 60/613,854
; PRIOR FILING DATE: 2004-09-28
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: Patent In Ver. 3.3
; SEQ ID NO 41
; LENGTH: 2710
; TYPE: PR1
; ORGANISM: Clostridium difficile
US-11-051-453-41
Query Match 4.1%; Score 269; DB 7; Length 2710;
Best Local Similarity 19.4%; Pred. No. 1.7e-06;
Matches 304; Conservative 257; Mismatches 499; Indels 506; Gaps 82;
QY 1 MXTNDIVTNNPNISLYKQSKDFIKKENINKKOPFILIKNKLFISIDNSTEANTESLL 60
DB 492 LKASDLIEFKFP-----ENNLSQLTEQEI---NSLWSFDQASAKYQFE--- 531
QY 61 KYIFEELNYSVEQKAGQIEGVESRVDIILLFEND--KDKASFNNKLKEAKKNNEPIPIED 118
DB 532 KYRVDYTGSLSED-----NGVDFNKTALDKNYLLNNKIPSNNVESAGSKY----- 579
QY 119 ILITIAEVKRPSTFDAKDKLKESEDQ---LYRYLNQYQKHGYILSGKVMRLYDKSKVL 174
DB 580 VHYIQLQGDIDSYEATCNLFSPKNPKNSIIIQNNMESAKSYFLSDG----- 627
QY 175 YGKRYVIEPNFKKIBEEKYKQEWVFLFYILIRKERYLKTNSVIEVEKEQIAKEKIIQ 234
DB 628 ---ESTILELNKYRIPERLKNKEK---VKVTFIGHGKDEFNTSEFARLSVDLSNEISSFL 681
QY 235 KTLKELTYERPDDSIYFVIAKNIYDKFKLS---GKEITQHTILASILEE----- 280
DB 682 DTIK--LDISPKRVENLIGCNMFSYDFNVETYPGK-----LLLSIMDKITSLPDVVK 734
QY 281 -SIITILRIFFIAYIEDNDIFKKILOENKLYRSSISFRYFFYDENTKKKLEYKKIITIFN 339
DB 735 NSIT-----IGANQYEVRLNSEGRKELLAHSGKWINKEEAIMSDLSKEYI--FPD 783
QY 340 LLDKGSDAIKFPVFNGLFSEDKVKVLLNNEGILLSIEIEILVKMLFFEBEKNIKDEKFKV 399
DB 784 SID-----NKLKAKSKNIPGLASISED-----IKTLLLDASVSPDTPKIL 823


```

Qy 472 SGAYTPDDLTDFMVISSEBOLKTKS--PLDIKIIDNSCSGSH-----FLISCLDYLTEK 525
Db 1000 NTG-ETKDFQKQWGIKDKQLEAKNTQTLTADVKDNEIISQQRKIFSLIQEKNELOQM 1058
Qy 526 VNYELDKFEDVKELDEEYRVIIIESEYD-VQDSISK--ELVLRKM-----LLKX 573
Db 1059 LESVIAEKQKLTDLKENIEMTENCQELRLGDELKQOEIVAOEKNHAIKKEGELSRT 1118
Qy 574 CIYGVDPINPISVEITMLSLWINTPIFGTPLSTIEHHIKAGNALLGYTKDEPFDIVKKPFE 633
Db 1119 C-----DRLAEVEEKLKESQSQLOKEKQOQOLLNVQEMSE 1152
Qy 634 SGFSLFKKRIKEITILEDIYOKIGINDTTKEIKSVIKEYEESKDIDNLIIFSL 693
Db 1153 -----MOKKINLENKNEKKNKELTLEHMETERLELAQKLNENYEEVAKSIKERVKLK- 1206
Qy 694 IKLYSLSPKSLNMFSDIASVISLIENILGNKNTSSEDK-----EKIEKIRKLSS 743
Db 1207 -----ELQKSPETERDHLRGYIREIE-ATGLQTKKEELKIAHHLKEHQETIDELRR--- 1256
Qy 744 YKFFHYGIEFDDIQGFDIVIGNPPWEK--TKPNETEPFSKH-----IPNVRKLG---- 792
Db 1257 -----SVSEKTAQIINTQDLEKSHTKLOE-EIPVLHBEQOELLPNKVEVSETQES 1303
Qy 793 -IKEQNIIOEILSKDNHPLS-LEYNEEKNSIIAINIYKF-----DFKFTSGGDPNLF 845
Db 1304 TMELELLTEQSTKDTTLARIEMER-----URLNE--KFOESQEIISLTKERD---- 1352
Qy 846 RYPTVFNLLKKE-----KGNLTVLVPSAIWNESSSRILRKHIFARYKLANIY 892
Db 1353 -----NLKTIKEALEVKHQDLKEHIRETLAKIQESQSQEQSLNKKEDONETTKIVSEM 1406
Qy 893 YQPENK-----KPKDVHSPFPAIFQLSNIKESTSPFAKMWIQSSDNI 937
Db 1407 EQPKPKDSALLRIEIMGLSKLOESHDEK-----SVAKEKDQLRLQELVQESDQ 1460
Qy 938 LKEITRD-----LKDSKODAYKGIELNINOIKK--LSPIQESIIIEF 976
Db 1461 LKENIKEIVANGHLETEELKVAHCCLEKEETIN-----ELRVNLSKEKETESTIITQKOLEAI 1517
Qy 977 KD-----NBEFTL--INKMFSKFSALGEGYIDFPKGLDPSIKNRKSLKKECNR 1022
Db 1518 NDQLQNKIQEIEKEQFNKQISEVQEKVNLKQ-FKEHRAKADQSALQIESKMLELN 1576
Qy 1023 KNILFLYSGANIHQNSRPFEDKAKESSKLMWDKEDLEKVUTKONQVTERVFRATA 1082
Db 1577 RL-----QESQEIQIMIKKEEMKRV--QALQIER---DQLK 1610
Qy 1083 SNTNERTMISTLSPGNCYCVNSIYINDEKTPISLYKKLFIISIFNSFPVDFLLRRFVDSN 1142
Db 1611 ENTK-----IVAKKESQEK-----YQFLQNTAVNET 1639
Qy 1143 VLKSLCYQCPMPQPEKEILSNPLYNLAKNTSLLIAKNDPENFKYLLYLEYKPKDKEV 1202
Db 1640 QERKCEIHLKEQETQKLNLENIENTIRLTQIL-----HENLEEMRSVTKERDRLRSV 1694
Qy 1203 NKILKLDKEPFEKKEKENENNFIASLYSLAKDFTLGLDFKALNKKKGGDIYISLLK 1262
Db 1695 EETLKVERD---QLKENRETITRLIKOLEKELKIV--HMHLKEHQETIDKLRGIVS 1748
Qy 1263 GYDNYLLN 1270
Db 1749 EKTNEISN 1756

```

RESULT 4

```

US-11-196-400-3
; Sequence 3, Application US/11196400
; Publication No. US20050287166A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES

```

```

; FILE REFERENCE: 200773USODIV
; CURRENT APPLICATION NUMBER: US/11/196,400
; CURRENT FILING DATE: 2005-08-04
; PRIOR APPLICATION NUMBER: US/09/742,096
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,462
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: P. falciparum
; US-11-196-400-3

Query Match 4.0%; Score 261.5; DB 7; Length 1786;
Best Local Similarity 20.0%; Pred. No. 2.5e-06;
Matches 285; Conservative 211; Mismatches 399; Indels 527; Gaps 71;

Qy 27 KENINKLDPFILIKNKLFSDIDNSTEA---NIESLLKYIFEEL-NYSVEQKQAGQI--- 79
Db 624 EENVEE-----SVAENVEESVAENVEESVAENVEEIVAPTVEEIVAPSVVES 678
Qy 80 -----EGVESRVDILLFENDKDKAFNNKLEAKKNNHEPIP--JEDIL--LIAEVKRT 129
Db 679 VAPSVESVEEENVEESVAENVEESVAENVEESVAENVEEIVAPTVEEIVAPTVEEIVAPT 738
Qy 130 FSPDAKDKLKESEDQLYRLNQYQKHGILSNGKWRMLYDKSKVLYGKRYIEFNPXKIE 189
Db 739 VEEIVAPSVVES-----VAPSVESVEEEN--VE 764
Qy 190 EKEEYKEQEWFLFYILIRKERYLKTNSVIEKEQIAKE-----KEIQTKLKEI 240
Db 765 ESAENVEE-----SVAENVEESVAENVEESVAENVEEIVAPTVEEIVAPSVVES 806
Qy 241 LYERPDSDSVFKIANNVD-----KEFKLS-----KREITQHLASILE----- 279
Db 807 VAPSVESVAENVATNLSNLLSNLGGIETEEIKDSILNEIEEVEKENVVTITILENVEET 866
Qy 280 --ESIIFILRIEPIAYIEDNDIFKKILOE--NKLYRSSISFRYFFFDENTKKKLEVKKII 335
Db 867 TAESVTTFSNI--LBEIQENTITNDTIEKULSELHENVLAAAL-----ENTQSEEEKKEVI 920
Qy 336 -----TIEN 339
Db 921 DVIEEYKEERVATTLIETVEQAEKSANTITEIPENLEENAVESNENVAENLEKLNETVFN 980
Qy 340 -LLDKGSDAKPP-----VFNGLFSE--DKVKYLNNEGLLS--ISIEEILLVWMLFF 387
Db 981 TVLVDKVEETVEISGESLENENDKAPFSEIFONVKGI-QENLLTGMFRSISIV--IQS 1037
Qy 388 EKNIKDEKFKVYSRLDPKSGFGLVETLLEYDLRIADT-----TVHRIIEDGYL----- 437
Db 1038 ESKVDLNNVVS-SILD--NIENMKGLNKLLENISSTEGVQETVTEHVEQNVDVDP 1094
Qy 438 -----IRTEEELENKKVVK-----IATYLK 457
Db 1095 AMKQDFLGLINEAGGLKEMFFNLEDVFKSESDVITVEEIKDPVQKEVEKETVSIIEEME 1154
Qy 458 GNIYLTSLDRKKSGAYTTPDDLTDFMVISSEBOLKTKSPLDIKIIDNSCSGSHFLIS 517
Db 1155 ENIV---DVLEBEK-----EDLTKM-IDAVEESIEISS----- 1184
Qy 518 CLDYLTKVWYELDKFEDVK--KELDEEYRVIIIESEYDQDSISKELVLRKMLLKXCI 575
Db 1185 -----DSKEETESIKDEKQVSLVVEEVQDNDMDSESVKVLKKNBEELMK 1231
Qy 576 YGVDPINPISVEITMLSLWINTPIFGTPLSFIEHHIKAGNALLGYTKDEPFDIVK-----K 630

```


QY 1035 HQFNSRFFE-----DKDAKESKLLWIDKEDLEKVLTKDNQYQTERVFRATA-SNTNE 1087
DB 1441 LQDEKRLERARIAQLEEELEEEQGNWMSDRVRKATQQAQESLNLATERSTAQKNESA 1500
QY 1088 RTWISTLSPGNCYCVNSIYINDKTPISLYKFLF-----IISIFNSFVDFLLRRFVDSN 1142
DB 1501 RQQL-----BRQNKELRSKLHEMGA VKSKFKSTI-----AALEAK 1536
QY 1143 VLKSCLYQCPMPQPEBEKEILSNPLYNLAKNTSLLAKNDPENFKYLLLYEYFKPKFKV 1202
DB 1537 IAQ--LEEQVEQEAQKATSKLSKQDKKLEILLQVEDERKM-----AEQYKEQAERG 1589
QY 1203 NKILKLDKEDPEPFKEKENENNFIASLYSLAKE-DFIT-----LLG-DFKALNNKKKG 1253
DB 1590 NARVKQLKQ--LEEAEEESQRINANRRKLQRLDEATESNEAMGREVNALSKLKG 1644

RESULT 6
US-10-995-561-661
; Sequence 661, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 661
; LENGTH: 1938
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-995-561-661

Query Match 3.48; Score 224; DB 6; Length 1938;
Best local similarity 19.9; Pred. No. 0.00024;
Matches 274; Conservative 211; Mismatches 544; Indels 348; Gaps 60;

QY 61 KYIFEELNYSVEQOKA-----GQEGVES-----RVDILLFENDKOKASFNKKLEA 107
DB 718 RIVQEFQRQRYEILANAIKPGFMDGQACILMIKALELDPNLRYICQSKIFRTGVLAH 777
QY 108 KQNEPIEDILIAEVKRPSTFDKOKLKSESDQ--YRYLNOYQKHGYILSNKQW 165
DB 778 LEEERDLKITDVIMAFQAM--CRGYLAKAFKRRQQQLTAMKVIQRNCAAYLKLNRWQW 835
QY 166 RLVDKSVLYGKRYIEFNFKEEYKEQEWFLFYILIRKERYLTSN---VIEVE 222
DB 836 RLFTVKRPL-----LQVTRQEEWQAKEDE-----LQTKERQOKAENELKELEQ 881
QY 223 KEQIAKEKEIIQKTL--KEILYERPDSDSIVFKAKN-----IYDKFKL----- 264
DB 882 HSQLTEKNLLOEQLOAETELYAEAEWVRLAAKQELSEILHEMARLEEEEDRCQQL 941
QY 265 --SGKETTHILASILEESIIFILRIFFIAYIEDNDIFKKILQBNKLYRSS-----ISFry 318
DB 942 QAERKQWQOQMLD--LEEQ-----LEEEAARQKQLEKVTAEAKIKLEBDEI 987
QY 319 FFYDENTKKLYKKII-----TIFNLLDKGSDAIKFPVNGGLFSQKVK---YLNNE 369
DB 988 LVMDQNNKLSKSKRLKEERISDITTTNLAEE-----EKKAKNLTCLKNK 1031
QY 370 GLLSISIEELIVQMLFEEKNIQD-EKFKVYSRLDPKSKFGEYVETLLEVDLRIADTTVH 428
DB 1032 HESMISLEVLKK-----EKSROELEKLRKLEGDADFE-----QIADLOA- 1076
QY 429 RIIBDGYVLINTEBELE-----NKKVNKIATYKNGIYTSRLDRKSKG-- 473
DB 1077 QIAELKQWAKKEELQAALRLDDEIATQKNALKKIRE-LEGHISDLQEDLDSEARA 1135

QY 474 -----AVYTPDDLTDPMWISSIEBQKTKSPLDIKIIDNSCGSHFLISCL 519
DB 1136 KAEKQKRLDGELEAKT--ELEDYLDSTATQELRAKREQEVTVL----- 1179
QY 520 DYLTEKWMYELDKFEDVKELDEEYRVYIEESEEVDQDSISK-----ELVXLMMLKK 573
DB 1180 -----KKALDEETR-----SHEAQVQEMRQKHAQVAVEELTEQLEQFQFR 1217
QY 574 CIYGVNDINPISVEITWLSLWINTFFIGTPLSFIEHHIKAGNALL-----GYTKDE 623
DB 1218 AKANLDKNQKQTTLEKNVADLAGELRVLGQAKQVEHKKKLEAQVOELQSKCSGDERARAE 1277
QY 624 FFDIVK--KPESGSLFKPKRIKEIITILEDIYQIKGINDT-----TKEDIKSKK 673
DB 1278 INDKVHKLQNEVESVTGMLNEAGKAIKAKOVASLSSQLQDTQELLQETQKLVNYSK 1337
QY 674 IYKYEESKDI-----DNLRITPSLKLKLSYSLFKSLNMFSDIASVISLIENI 722
DB 1338 LRQLEERNLQDLDEEMAKQNLERHISTLNI-QLSDSKK---KLQDFASTVEALEE- 1392
QY 723 LGNKTSSSEDEKI-----EKIRKLSYYKFFHYGIEFPDIOEGFDIV----- 764
DB 1393 -GKGRFQKEIENLTQOYEKAAAYDKLEKTKRLQ-----QELDDLVDLDNQ 1440
QY 765 --IGNPPWEKTKFNETEPFSKHIPYVRKLGIKIKEONIIKOEILSKDNHPLSI-----EYNE 817
DB 1441 QLVSNLEKQKQKFDQLLAEEKNISKY---ADERDRAEAREKETKALSLARALEALE 1497
QY 818 EKNSIIAINNIYKFDPKFTSGDPNLFYFVTFNL-KLIKKGNI--TYLVPSAINNESS 875
DB 1498 AXEELERTNKLKAEMLDVSSKDD-----VGKNVHELEKSKRALEQTQMEEMKTQLEEL 1551
QY 876 SRILRKHIFARYKLVYI-----OPENKRFKDVHSSFKFAIFQLSNIKESISSFKAKFM 930
DB 1552 EDELOATEDAKLRLEVNMQALKGQPERDLQARDEQNEEK-----RRQLQRLHEYTELE 1606
QY 931 IOSSDNIL-----KEITRDLK--SKDAYKGIELNINOIKGL-SPIOESIIEPKD--- 978
DB 1607 DERKQALAAAKKLEGLDKLELQADSAIKREAIKQLRKLQAKMDQFQLEEDARA 1666
QY 979 --NEEFTLKNKFSKPSALGEGYIDPKGLDPSIKNRK--SLLKECNKNILFLYSGANI 1034
DB 1667 SRDEIFATAKENEKKALEADLMQLEDAARAKQADLEKEELAEELASSLSGRNA 1726
QY 1035 HQFNSRFFE-----DKDAKESKLLWIDKEDLEKVLTKDNQYQTERVFRATA-SNTNE 1087
DB 1727 LQDEKRLERARIAQLEEELEEEQGNWMSDRVRKATQQAQESLNLATERSTAQKNESA 1786
QY 1088 RTWISTLSPGNCYCVNSIYINDKTPISLYKFLF-----IISIFNSFVDFLLRRFVDSN 1142
DB 1787 RQQL-----ERQNKELRSKLHEMGA VKSKFKSTI-----AALEAK 1822
QY 1143 VLKSCLYQCPMPQPEBEKEILSNPLYNLAKNTSLLAKNDPENFKYLLLYEYFKPKFKV 1202
DB 1823 IAQ--LEEQVEQEAQKATSKLSKQDKKLEILLQVEDERKM-----AEQYKEQAERG 1875
QY 1203 NKILKLDKEDPEPFKEKENENNFIASLYSLAKE-DFIT-----LLG-DFKALNNKKKG 1253
DB 1876 NARVKQLKQ--LEEAEEESQRINANRRKLQRLDEATESNEAMGREVNALSKLKG 1930

RESULT 7
US-10-995-561-662
; Sequence 662, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561


```
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 662
; LENGTH: 1938
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-993-561-662

Query Match      3.4%; Score 224; DB 6; Length 1938;
Best Local Similarity 19.9%; Pred. No. 0.00024;
Matches 274; Conservative 211; Mismatches 544; Indels 348; Gaps 60;

QY 61 KYIFEELNVSQOKA-----GOIEGVES-----RVDILLFENDKDKASFNKLEA 107
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 718 RIVQEFQRORYEILAAIPKFGMDGQACILMIKALELDPNLYRIGOSKIFFTGVLAH 777
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 108 KNNNEPIPIEDILLIAEVKRPFPSPDAKDKLSESDOL--YRYLNQYQKHGILSNGKYM 165
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 778 LEEERDLKITVIMAFQAM--CRGYLARKAFAGROOQLTAMKVIQRNCAAYLKLRNWM 835

QY 166 RLYDSKVLGYEKRYIEFNFXKIEEKEVKEBWFVLYLIRKERYLXTSN---VIEVE 222
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 836 RLFTVKVPL-----LQVTQEEEMQAKEDB-----LQTKERQQAENELKELOK 881

QY 223 KEQIAKEKEIIQKTL--KEILYERPDSDSIVPKIAKN-----IYDKFPL----- 264
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 882 HSQLTEKNLQEOQLQAEYELVAEEMVRVLAACKQOELEELHMEARLEEBEDRGQOL 941

QY 265 --SGKEITOHILASILEESIIFLIRFFIATYEDNDIFKKILOENKLYRSS-----ISFVY 318
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 942 QAEKQKMAQOMLD--LEEQ-----LEEEAARQKLOLEKVTAABAKIKKLEDEI 987

QY 319 FFYDENTKKLEVKKII-----TIFNLLDKGSDAIKFFVFNGLFSEDKVK---YLNNE 369
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 988 LVWDDQNKLSKERKLEERISDLTNTLAE-----BEKAKNLTCLKNK 1031

QY 370 GLLSISEIBEILVKMLPFEEKNIKD--EKVYKSRDPKSGELYETLYEYDLRIADTVH 428
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1032 HESMISELEVLK-----EKSQLEKLEKRLKLEGASDFHE-----QIADLQA- 1076

QY 429 RIIEGCVYLIRTEEBLE-----NKKVNIATYKGNIVLTSLRDRKSG-- 473
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1077 QIAELKQWLAKEEELQALARLDDEIAQKNALKKIRE--LEHISDLQEDLDSEARA 1135

QY 474 -----AYTPDLDLDFWISSIEEQLATKSPDKIKI1DNSGSGHFLISCL 519
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1136 KAKQKRDLGSELEALKT--ELEDTLDDSTATQELRAKREQEVTVL----- 1179

QY 520 DYLTEKVVWYELDKFEDVKKELDEYRVITIESEYDVQDSISK-----ELVLKRMLLKX 573
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1180 -----KALDEBTR-----SHEAQVQEMRQGAQVAEELTQLEQPKR 1217

QY 574 CYIGVDINPISVITWLSWINTPIFGTPLSFTEHHIKAGNALL-----GYTKOE 623
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1218 AKANLDKNQTLKEKNADLAGELRVLQQAQVEHEKKKLEAQVQELQSCSDGERARAE 1277

QY 624 FPDIVVK--KPSGSLFKRKRIKEITILEDIYQIKGINDP-----TKEDIEKSK 673
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1278 LNDKVHKLQNEVESVTGMLNABEGKAIKAKOVASLSQLOQDQELLQEBETROKJNVSTK 1337

QY 674 IYKEVEESKDI-----DNLRIFSLIKYLSLSPDKSLNMFSDIASVISLIENI 722
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1338 LRQLEERNSLOQLDEEMAKQNLREHISTINI--QLSDSK--KLQDFASIVEALEE- 1392

QY 723 LGNKTSSSEDKBI-----EKIKLSSYYKFFHYGIEPDPDIOEGFDIV----- 764
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1393 -GKKRFQKEIENLTQVVEKAAAYDKLETKNRLQ-----QELDDLVDLDNQR 1440

QY 765 --IGNPPWEKTKNETETFFSKHPIPNYKLGIKIKNQIILKQILSKDNHPLSI-----EYNE 817
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1441 QLVSNLEKQKQFDQLAEBEKNISSKY---ADERDRAEABAREKETKALSRLALEALE 1497
```

RESULT 8

```
US-11-077-550-20
; Sequence 20, Application US/11077550
; Publication No. US2005024435A1.
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; PRIOR FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 1169
; TYPE: PRF
; ORGANISM: Clostridium botulinum
US-11-077-550-20
```

```
Query Match      3.4%; Score 222; DB 7; Length 1169;
Best Local Similarity 20.0%; Pred. No. 0.00017;
Matches 196; Conservative 169; Mismatches 306; Indels 310; Gaps 51;
```

```

QY 11 NPWISLVKQSKD-----FIKKENINKL-----KDF 36
Db 314 NININIVKFKDKYFVEDSEKYSIDVSPDLKYLKSLMFGFTETNIAENYKIKTRASY 373
QY 37 F-----ILFKNKL-----FSIDNSTEANI-----BSLKUYFEELN-----YS 70
Db 374 FSDSLPVKIKNLDNEIYIIEBGFNISDKMEKEYRGONKAIKQAYEISEHSLAVYK 433
QY 71 VEOQKAGQIBGVSRVD--ILLFENDK----DKASFNNKLKEAKKN--EPIPIEDIL- 120
Db 434 IOMCKSVKAPGICIDVDNEDLFIADKNPSFSDLSKNERIEYNTQSNYIENDFPINELIL 493
QY 121 ---IIAEVKRPT-----FSPD-----AKDKLKBSEDQLYVLYNQYQKHGCI-LS 160
Db 494 DTLILSKIELPSENTSLTDFNDVDPVYVEKQPAIKKIFTDENTIFQL--YSQTFPDIR 551
QY 161 NGKWRLYDKSVLYGKRYIEFNFXKIEEYKEQWFLVFIYLIRKERYLKTNSVIE 220
Db 552 DISLTSSFDDA-LIFSNNKVSFFSMDYIKTANKVVEAGLPAGWV-----KQIVNDFVIE 604
QY 221 VEQOTAKEKEIIQKTLKEILYERPDSDIVF-----KIANIYDKEPKLSGKEI 269
Db 605 ANK-----SNTMDKI-----ADSLIVPYIGLALNVGNETARGNFENAEFIAGAS- 649
QY 270 TQHLASILEESIIFILIRIFFI-AYIED-NDIFKKILQENKL-----KIANIYDKEPKLSGKEI 269
Db 650 ---ILLEFPILLIPVVGAFLESYIDNKNKIITKI--DNALTKRNEKSDMYGLLVAQW 704
QY 311 RSSISRYFPYDENTKKLEY-----KKII-----TIFNLDDKGSDAIKFPVFNGLFSED 361
Db 705 LSTVNTQFYTIKSGMYKALYQAALAEELIKRYVNIYSEKEKSNINIDFNDINSKL---- 760
QY 362 KVKYLANNEGL-LSISEIEELVKMLFPEEKNIIDKFPVKYSRLDPKSGFELYETLLEYDL 420
Db 761 -----NEGNOAIDNNNPI-----NGCSVSYLMKKMIPLAVEKLLDFDN 800
QY 421 RIADTTVHRIEDGVYLRTEEBELNKKVNKIATYLGKNIYLSRSLDRKSKSAYTTPD 480
Db 801 TLKNLLNYIDENKLYLIGS-AEYKSKVN---YLK-----TIMPFDL-----SITYNDT 847
QY 481 LTDFMVISSIEEQLKTSPLDIKIIDNSCGSGHFLISCLDYLTKVWYELDKPEDVKKEL 540
Db 848 IL-LEMPKNYSSEILANNIILNRYKDN-----LIDLSCY-----GAKVEYVDGV--EL 893
QY 541 DEEYRVIIIESEYVDQDSISKELVLKMLLKXICIVGVDINPISVEITMLSLMINTFIIG 600
Db 894 NDKNQFKLTSSANSKIRVTQNIIFNSVFLDFSV-----SFWIRIPKY- 937
QY 601 TPLSFIEHHIKAGNALLGYTKDFPDIVKKKFSGSGSLFKKRIKEIITILEDIYQKIKGI 660
Db 938 -----KNDGQNYIHNETHINCWNNSGWSIRGNRIIWLID-----ING- 980
QY 661 NDTTKEDIEKSKIKIYEYESKDIDNU--RIIFSLI-----KLXSLSPDKSLNMEFSD 711
Db 981 -----NTKSVFFPEYNIREDISEYINRWFFVTITNLLNNAKIY-INGKLESNTDIKD 1030
QY 712 IASVISLIE-----NILGNKTSSEDEKIEKIRKLSYKFFHYG 751
Db 1031 IREVIANGETIIFKLDGDIRTOFTWVKYFSIFNTLSQSNIEERYKIQSYSEYKDF-WG 1089
QY 752 IEPDIDEGPDIVGN-----PPWE---KTKFNETERFESHIPNRYKLGIBKQN 797
Db 1090 NPLMYNKEYTWFNAGNNSVILKUDSPVGEILLTRSKYNQ-----SKYI-NYRDLYIGERF 1145
QY 798 IIKQ-----EILSKDNH 809
Db 1146 IIRKNSQSINDDIVKEDY 1166

```

RESULT 9
US-10-995-561-660
; Sequence 660, Application US/10995561

```

; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 660
; LENGTH: 1954
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-995-561-660

Query Match 3.4%; Score 219.5; DB 6; Length 1954;
Best Local Similarity 19.9%; Pred. No. 0.00041;
Matches 274; Conservative 213; Mismatches 544; Indels 349; Gaps 61;

QY 61 KVIPELNVSVQOQA-----QOIEGVES-----RVDILLFENDKDKASPNKLEA 107
Db 718 RVFOEQRORYEILAAANAIPKGMGKQACILMIKALELDPNLYRIGOSKIPFTGTGLAH 777
QY 108 KKNNEPIPIEDILIIAEVKRPTFSFDADKDKLKESEDQL--YRYLNOYQKHGILSGKW 165
Db 778 LBEERDLKITDVIAPQAM--CRGYLAKAFAPAKRQOQLTAMKVIQRNCAAYLKLWQW 835
QY 166 RLYDKSVLYGKRYIEFNFXKIEEYKEQWFLVFIYLIRKERYLKTNS---VIEVE 222
Db 836 RLFTKVKPL-----LQVTRQEEMQAKED-----LQTKERQQAENELKELEQ 881
QY 223 KQOIAKEKEIIQKTL--KEILYERPDSDIVFKIKN-----IYDKFKPL----- 264
Db 882 HSQLTEERKNLQEQOQAETELVAAEEMVRVLAACKQOELEELIHEARLEEBEEDRGOQL 941
QY 265 --SGKEITOHILASILEESIIFILIRFFIAYIEDNDIFKKILOENKLYRSS---ISFRY 318
Db 942 QAEKRWAAQOMLD--LEEQ-----LEEEAARQKQLEKVTABAKIKKLEDEI 987
QY 319 FFYDENTKKLEYKIKII-----TIFNLDDKGSDAIKFPVFNGLFSEDVKV---YLNNE 369
Db 988 LWMDQNNKLKSKERKLEBERISDLITNLAEE-----BEKAKNLTKLKNK 1031
QY 370 GLLTSISEIEELVKMLFPEEKNIKD-EKFPVKYSRLDPKSGFELYETLLEYDLRIADTTVH 428
Db 1032 HESMISELEVRLKK---BEKSRQLEKLRKLEGDASDFHE-----QIADLQA- 1076
QY 429 RIIEEDGVYLRTEEBELN-----NKKVNKIATYLGKNIYLSRSLDRKSG-- 473
Db 1077 QIAELKQAKKEEELQAAALARDDEIAQNNALKKIRE-LEGHIISDLQEDDLSEARA 1135
QY 474 -----AYYTPDDLTDFMVISSIEEQLKTSPLDIKIIDNSCGSGHFLISCL 519
Db 1136 KAEKQKRDIGEELEALKT--ELEDTLDDSTATQELRAKREQEVTVL----- 1179
QY 520 DYLTEKWYELDKFEDVKELDEEYRVIIIESEYVDQDSISK-----ELVLKRMLLKX 573
Db 1180 -----KKALDEETR-----SHEAQVQEMRQKHAQAVEELTEQLQEPKR 1217
QY 574 CIYGVDDINPISVEITMLSLMINTFFIRGTPLSPTSEHHI KAGNALL-----GYTKDE 623
Db 1218 AKANLKNKQITLEKENADLAGELRVLGQAQOEVEHKKKKLEAQVQELQSKCSGERARAE 1277
QY 624 FFDIVKK---KPFESGSLFPKGIKEIITILEDIYQKIKGINDT-----TKEDIEKSK 673
Db 1278 LNDKVHKLQNEVESVTGMLNEAEGKAIKLAKOVASLSQLODTQELLQEBETROKLVNSTK 1337
QY 674 IYKEYEESKDI-----DNLRIFIISLKYLSYSPDKSLNMEFSDIASVISLIENI 722
Db 1338 LRQLEERNSLQDLDEEMAKONTERHISTLNI-QLSDSKK---KLQDPASTVEALEE- 1392

```

QY 723 LGNTSSDEKKEI-----EKIRKLSYYKFFHYGIBFPDIOEGFDIV-----764
DB 1393 -GKRFQKEIENLTQOYEBKAAAYDKLETKRNLQ-----QELDDLVDLDNR 1440
QY 765 --IGNPPWEKTFNTEFFSKHPIPNYKLGIGKEONIIKOBILSKDNHPLSI-----EYNE 817
DB 1441 QLVSNLEKKQKQKFDQLLAEBEKNISSKY---ADERDRAEAAREKETKALSARALEALE 1497
QY 818 EKNSIIAINNIYKFDKPTSGDGNLFRYFTFNL-KLIKEKGNL-TYLVPSAINNESS 875
DB 1498 AKEELERTNMLKAEMEDLVSSKDD-----VGNVHELEKSKRALETQMEEMKTQLEEL 1551
QY 876 SRLRKHIPIARYKLANIY-----OPENKKRFDKDVHSSPKFAIFQLSNISTESSPKAKFM 930
DB 1552 EDELQATEDAKLRLEVNMQALKGQFERDQARDEQNEEK-----RRQLQRLHEYTELE 1606
QY 931 IOSSDNIL-----KEITRDLK---SKDAYKGIELNINOIKKL-SPIQESIIEFKD---978
DB 1607 DERKQALAAAANKKLEGLDKOLEQADSAIKGREBAIKQLRKLQAKMDQFQLEEDARA 1666
QY 979 --NEEPTLINKMFSKFSALGEGYIDFKGLDPSIKNRK--SLLKECNKNLPILYSGANI 1034
DB 1667 SRDEIFATAKENEKAKSLEADLMQLQEDLAAARARKQADLEKEBELABELASSLSGRNA 1726
QY 1035 HQNRSRFPF-----DKDAKESKLLWIDKEDLVKLTQDNQYQTVFRVRAIA-SNTNE 1087
DB 1727 LQDEKRLERARIAQLEEELEEEQNGNMEASDRVRKATQAEQLSNELATERSTAQKNESA 1786
QY 1088 RTWISTLSPGNCYCVNSIYINDEKTPISLYKKLF-----LISFNSFVDFLLRRFVDSN 1142
DB 1787 ROOL-----BRQNELRSKLHEMEGAVKFKSTI-----AALEAK 1822
QY 1143 VLKSCILYQCPMPQPEEKEILSNPLYNLAKNTSLIAXNDPENFKYLLYLEVPKPKBKV 1202
DB 1823 IAQ--LEEQVEQAREKQATKSLKQDKKXKLEILAQVEDERKM-----AEQYKEQAEG 1875
QY 1203 NKILKLDKDEDFYKVENENNFIASLYSLAKB-DFIT-----LLG-DPKALKNK-KKGED 1255
DB 1876 NARVKQLKRGQ--LEEABESQRINANRRKLQRELBATEESNEAMGVEVNAKSLKLRGNE 1933

RESULT 10

US-10-995-561-664
; Sequence 664, Application US/10995561
; Publication No. US20050272054A1

GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 664
; LENGTH: 1972
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-664

Query Match 3.4%; Score 219.5; DB 6; Length 1972;

Best Local Similarity 19.9%; Pred. No. 0.00042;
Matches 274; Conservative 213; Mismatches 544; Indels 349; Gaps 61;

QY 61 KYIFEELNYSVEQOKA-----GOIEGVES-----RVDILLFENDKDKASFNKKLEA 107

DB 718 RIVQEFQRVYEILAAAIKPGFMDGQACILMIKALELDNLYRIGQSIFPRTGVLAH 777

QY 108 KQNEPIPIEDILLIAVRKPTSPFADKDKLESEDLQ--YRYLNQYQKHGILSKGW 165

DB 778 LEEERDLKITDVIMAFQAM--CRGYLARKAFAKRQOQLTAMKYIQRNCAAYLKLNRQW 835

QY 166 RLYDKSKVLYGKRYIEFNFXKIEBEKYEKEQEWVLFYILIRKERYLKTSN---VIEVE 222
DB 836 RLFTVKFPL-----LOVTRQEEEMQAKDE-----LOKTKERQQAENELKELEOK 881
QY 223 KEQIAKEKEIIOKTL--KEILYERPDSDIVFKIAKN-----IYDKEFKL-----264
DB 882 HSQLETKENNLQEQQAETELVAAEEMRVRLAAKQOELEELHEMEARLEEBEDRGQQL 941
QY 265 --SGKEITQHILIASILERSIIFILRIFFIAYIEDNDIPKKILOENKLYRSS-----ISFRY 318
DB 942 QAEKCKMAQMLD--LEEQ-----LEEBEAAARQLEKVEKVTAAAKIKKLEDR1 987
QY 319 FPDYDENTYKCLYKKI1-----TIFNLLDKGSDAIKPFVNGFLFSEDKVK--YLNNE 369
DB 988 LVMDQNNKLSKERKLEERISDLTNTLAE-----BEKANNTKLKRNK 1031
QY 370 GLLSISEIEEILVQMLFFEEKNIKD-EKPFVKYSRLDPKSFGBELYETLLEYDRIADTVH 428
DB 1032 HESMISELEVLK---BEKSRQBLEKLRKLEGDASDFHE-----QIADLQA- 1076
QY 429 RIIEPDGVYLIRTEEBE-----NKKVNKIATYLGKNIYLTSRSLDRKKS---473
DB 1077 QIAELKQMLAKKEBELQALARLDDEIAQKNNAKKIRE-LEGHISDLQEDLDSEARA 1135
QY 474 -----AYYTPDDLTDPMVVISSIEEQLKTKSPDIKIIDNSCGSGHFLISCL 519
DB 1136 KAEKQKRDIGEELEAKT--ELEDLTDSTATOQELRAKEEQEVTVL-----1179
QY 520 DYLTEKVMYELDKFEDVKKELDEEYRVIIIEBSEEDVDQDSISK-----ELVKRMLLKK 573
DB 1180 -----KKALDEETR-----SHEAQVQEMRQKHQAQVLELTOEQFQR 1217
QY 574 CIYGVDIINPISVEITMLSIWINTFIFGTPLSPIEHKIKAGNALL-----GYTKDE 623
DB 1218 AKANLDKNKOTLEKENADLAGELRVLQAKQVEHKKKLEAQVQESQKSCSDGERARAE 1277
QY 624 FFDIVK---KPESEGSFLPKKRIKILITILEDIYQIKGINDT-----TKEDIKSK 673
DB 1278 LNDKRVHKLQNEVSTGMLNEAGKAIKLAQDVASLSSQLQDTQELLQEBETQKLVNSTK 1337
QY 674 IYKVEESKDI-----DNLRIIFSLIKLYSLSPDKSLNMBFSDIASVISIENI 722
DB 1338 LRQLEEEERNLQDQDEMEAKQNLERHISTANI-QLSDSKK---KLQDFASTVEALER- 1392
QY 723 LGNTSSDEKKEI-----EKIRKLSYYKFFHYGIBFPDIOEGFDIV-----764
DB 1393 -GKRFQKEIENLTQOYEBKAAAYDKLETKRNLQ-----QELDDLVDLDNR 1440
QY 765 --IGNPPWEKTFNTEFFSKHPIPNYKLGIGKEONIIKOBILSKDNHPLSI-----EYNE 817
DB 1441 QLVSNLEKKQKQKFDQLLAEBEKNISSKY---ADERDRAEAAREKETKALSARALEALE 1497
QY 818 EKNSIIAINNIYKFDKPTSGDGNLFRYFTFNL-KLIKEKGNL-TYLVPSAINNESS 875
DB 1498 AKEELERTNMLKAEMEDLVSSKDD-----VGNVHELEKSKRALETQMEEMKTQLEEL 1551
QY 876 SRLRKHIPIARYKLANIY-----OPENKKRFDKDVHSSPKFAIFQLSNISTESSPKAKFM 930
DB 1552 EDELQATEDAKLRLEVNMQALKGQFERDQARDEQNEEK-----RRQLQRLHEYTELE 1606
QY 931 IOSSDNIL-----KEITRDLK---SKDAYKGIELNINOIKKL-SPIQESIIEFKD---978
DB 1607 DERKQALAAAANKKLEGLDKOLEQADSAIKGREBAIKQLRKLQAKMDQFQLEEDARA 1666
QY 979 --NEEPTLINKMFSKFSALGEGYIDFKGLDPSIKNRK--SLLKECNKNLPILYSGANI 1034
DB 1667 SRDEIFATAKENEKAKSLEADLMQLQEDLAAARARKQADLEKEBELABELASSLSGRNA 1726
QY 1035 HQNRSRFPF-----DKDAKESKLLWIDKEDLVKLTQDNQYQTVFRVRAIA-SNTNE 1087
DB 1727 LQDEKRLERARIAQLEEELEEEQNGNMEASDRVRKATQAEQLSNELATERSTAQKNESA 1786
QY 1088 RTWISTLSPGNCYCVNSIYINDEKTPISLYKKLF-----LISFNSFVDFLLRRFVDSN 1142

```
Db 1787 RQQL-----ERONKELRSKLHEMGA VKSKFKSTI-----AALEAK 1822
QY 1143 VLKSCLYQCPMPQPEBEKEILSNPLYNLAKNTSLTAKNDPENFKYLLYLEVPFKDKVK 1202
Db 1823 IAQ--LEEQVEQBEAREKQAATSKLQKQKKEILLQVEDERKM-----AEQYKEQAEKG 1875
QY 1203 NKLKLDKDEPFKEKENENFIASLYSLAKE-DFIT---LLG-DFKALKNK-KKGRD 1255
Db 1876 NARVKQLKQK--LEEAEEESORINANRRKLQRELDATESNEAMGREVNALSKLRGNE 1933

RESULT 11
US-10-995-561-666
; Sequence 666, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 666
; LENGTH: 1972
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-666

Query Match 3.4%; Score 219.5; DB 6; Length 1972;
Best Local Similarity 19.9%; Pred. No. 0.00042;
Matches 274; Conservative 213; Mismatches 544; Indels 349; Gaps 61;

QY 61 KYIFEELNYSVEQOKA-----GOIEGVES-----RVDILLFENDKOKASPNKKLEA 107
Db 718 RIVQEFQRYEILANAIPKGMWGQACILMIKALEDPNLYRIGOSKIFPRTGVLAH 777
QY 108 KQNEPIPIEDILIIAEVRKPTSPFADKDKLKESEDL--YRYLNOYQKHGYLSNGKVM 165
Db 778 LEEERLDKITDVIMAFQAM--CRGYLARKAFARQQOQLTAMKVIQRNCAAYLKLNRNQWM 835
QY 166 RLYDKSKVLYGEKRYIEFNFKLKEEYKQEWFLFYILIRKERYLTSN---VIEVE 222
Db 836 RLFTKVKPL-----LQVTROEEEMQAKEDE-----LQKTKERQQAENELKELEQ 891
QY 223 KEQIAKEKEIIQKTL--KEILYERPDOSIVFKIAKN-----IYDKFEKL----- 264
Db 882 HSQLTEKNLLQEQLOAETELYABAEEMRVLAACKQOEELIHEMARLEEBEDRQOOL 941
QY 265 --SGKEITQHILASILEESIIFILRIFPIAYIEDNDIFKKILQENKLYRSS-----ISPRY 318
Db 942 QAERKQMAQQMLD--LEEQ-----LEEBAARQKLQLEKVTAEBAKIKLEDEI 987
QY 319 PFYDENTKKKLEVKYKII-----TIFNLLDKGSDAIFPVFNGGLSPEDKVK---YINNE 369
Db 988 LVNQDQNNKLKSKERKLEERISDLTNTLAE-----EEKAKNLTKLK 1031
QY 370 GLLSISIEBILVKMLPFPEKNTKD-EKFVKYSLRDPKSGVELVETLLEVDLRIADTVH 428
Db 1032 HESMISELEVRLLK-----EKSQOELEKLRKLEGDASDFHE-----QIADLQA- 1076
QY 429 RIIEDGVYLIRTEEBEEL-----NKKVNKIATYALKGNIYLTSLDRKSG-- 473
Db 1077 QIAELKQWLAKKEEELQAALRLDDETAQKNNAKKIRE-LEGHISDLQEDLDSERAARN 1135
QY 474 -----AYYTPDDLDFMVISIEBOLKTKSPDLTKIINDSCSGHFLISCL 519
Db 1136 KAEKQKRDLEGELEALKT--ELEDTLSDTATQQLFRAKREQEVTVL----- 1179
QY 520 DYLTEKVMYELDKPVEDVKELDEYRVIIIESEBEYVDQDSISK-----ELVLKRLMLKX 573
```

```
Db 1180 -----KKALDEETR-----SHEAQVQEMRQKHAQAVEELTEQLQFQFKR 1217
QY 574 CIYGVNDINPISVEITMWSLWNTTFCTPPLSFIEHHIKAGNALL-----GYTKOE 623
Db 1218 AKANLDKNQKTLTEKENADLAGELRVLGOAQEQVEHKKKLEAQVQLOSKCSDGERARAE 1277
QY 624 PFDIVYK--XPESGFSJFKKRIEIIITLIEDIYQKIKGINDT-----TKEDIEKSKK 673
Db 1278 LNDKVHKLQNEVESVTGMLNEAGKAIKLAKDVASLSQLODTQELLQETQKLNVSFK 1337
QY 674 IYKYEESKDI-----DNLRIIFSLIKYLSLSPDKSLNMEFSDIASVISLIENI 722
Db 1338 LQALEEERNSLQDLDEEMEAQNLERHISTLNI-QLSDSKK--KLQDFASTVEALEE- 1392
QY 723 LGNTSSSEDKKI-----EKIRLSYYFFHYFFGIEFPDIOEGFDIV----- 764
Db 1393 -GKGRFOKEIENLTQOYEEKAAAYDKLEKTKNRLQ-----QELDDLIVVDLNDOR 1440
QY 765 --IGNPPWEKTKFNETEPFSKHPNVRKLGIKEOIILKOEILSKDNHPLSI-----EYNE 817
Db 1441 QUVSNLEKKQKRFQDQLAEBEKNISSKY--ADDERAAEAREKETKALSLARALEEAL 1497
QY 818 EKNSIIAINNIYKDFCKFTSGDPNLFYFVTFNL-KLIRKGNL-TYLVPSAIWNSS 875
Db 1498 AXEELERTNKMLKAEMEDLVSSKDD-----VGKNVHELEKSKRALETQWEMKMTQLEEL 1551
QY 876 SRILRKHIFARYKLYNIY-----QFENKGRFKDVHSSFKFAIFQLSNIXESTSSFKAKFM 930
Db 1552 EDELQATEDAKLRLEVNMQALKGQFERDLQARDEQNEEK-----RRQLQRLHEYTELE 1606
QY 931 IQSSDNIL-----KEITRDLKD--SKDDAYKGIELNINOIKKL-SPIQESIIEPKD--- 978
Db 1607 DERKQALAAAKKLEGLKLELDLEQADSAILGREAIKQLRKLQAKMDQFORELDARA 1666
QY 979 --NEEFTLINMFKSPSALGEGYIDFKGLDPSIKNRK--SLLEKCNKNKLIPLYSGANI 1034
Db 1667 SRDEIFATAKENEKKAASLEADLMQLQEDLAAERARKQADLEKELEAEELASSLSGRNA 1726
QY 1035 HGFNRRFFE-----DKAKESSKLLWIDKELEKVLTKDNQYQTVRYRAIA-SNTNE 1087
Db 1727 LQDEKRLRLARIAQLEEELEEBEQNGMEAMSDRVKATQAQAEQLSNELATERSTAQKNSA 1786
QY 1088 RTMISTLSPGNCVYNSIYINDEKTPISLYKKLF-----IISFNSFVDFELRRFVDSN 1142
Db 1787 RQQL-----ERONKELRSKLHEMGA VKSKFKSTI-----AALEAK 1822
QY 1143 VLKSCLYQCPMPQPEBEKEILSNPLYNLAKNTSLTAKNDPENFKYLLYLEVPFKDKVK 1202
Db 1823 IAQ--LEEQVEQBEAREKQAATSKLQKQKKEILLQVEDERKM-----AEQYKEQAEKG 1875
QY 1203 NKLKLDKDEPFKEKENENFIASLYSLAKE-DFIT---LLG-DFKALKNK-KKGRD 1255
Db 1876 NARVKQLKQK--LEEAEEESORINANRRKLQRELDATESNEAMGREVNALSKLRGNE 1933

RESULT 12
US-10-793-626-2964
; Sequence 2964, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P034800S
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2964
; LENGTH: 5024
; TYPE: PRT
```


Db 226 VOTLMDNMTLERERESEKLDCKQLSQSVASSEATISQLRSELAKGPQEVAVVYQELQK 285
QY 44 LPSIDDNSTAN---TESLLKYIFEELNYSVEQKAGQEGVESRVDILLFENDKDKASF 100
Db 286 LKSVNELTQKQNTLTENLLK---KEQDYTKLEERINEESVSKNIQATLHQKDLDCQQL 342
QY 101 NNKLKBAKKNNEPIEDILLIAEVRKPTFFSDAKDKLJESDQL---YRYL----- 149
Db 343 QSLASASETSLRHIVB---LSE-----KGEATQKLAEELSEVETKYQHLKAEFPKQLQ 392
QY 150 --NOYQKHGYILNGKWRLYDKSVLYGKRYIEFNFXXKIBPKBY-----KEQWFP 200
Db 393 QOREEKEQHLQQLQSEINQLH---SKLLETERQLGEAH---GRLEKQROLSEKLMKDEKQOVA 449
QY 201 VLFYILIRKERYL---KTSNVIEVEKE-----QIAKEKELIOKTLKEILYERPDSDIVFKI 253
Db 450 DLQKLRLUEQKEKVTNSTELQHQDLTKYQHQEQOALQSQSTTAKLREAGND---LEQV 507
QY 254 AKNIYKFKLSGKEITOHILASILBESIIFILRIFFIAYIEDNDIFKKI----- 303
Db 508 LRQIGDKQKIQNLE---ALLQKSENISLLEK-----EREDLYAKIQAGEGETAVL 556
QY 304 -----LQE-----NKLYRSSISFRYFFVDENNTKKLYKKIITIFNLDKGSDAI 348
Db 557 NQLQEKNNHTLQEQVTQLTEKLKNQSESHKQ---AQENLHDQVQSKA-----HLRAAQDRV 609
QY 349 KPFVFNGLFSEDKVXYLN---NEGILLSISEIIEILVXMLFPFEKNIKDEKFKVYKSRLOP 405
Db 610 -----LSLETSVNELNSQMSKEKVSQQL-DIQIK-----AKTELLLS----- 646
QY 406 KSPGELYETLELDLADTTFVHRIEDGVYILRTBELEENKKNKIATVYKGNIVYLSR 465
Db 647 ---AEAATAQRADLQNLHTAQNALQD-----KQBELNKITQLD---QVTAK 689
QY 466 SLDRKXGAYTTPDDLTDFWVSIIEBQTKTSPLDIKIDNSCGSHFLISCLDYLTREX 525
Db 690 LQDKQEH-----CSQLESHLKEYKEKYL-----LEQKTEE 720
QY 526 VWTELQKFEVDKELDEYRVIIESEEDVDQD-----SISKELVLRM 569
Db 721 LEGQIKKLEADSLE-----VKASKEQALQDLQOQROLNTDLRLATLSKQLEMEKE 772
QY 570 LLKXCIVGVDPINPISVEITWLSLWINTFIPGTPLSPTEHHIKAGNALLGYTKDEFEDIVK 629
Db 773 I-----VSTFLDL-----Q 782
QY 630 KPFEQSLPKRKIKSIIITILEDIYOKIGINDTTKEDIKSKKIYKVEES----- 681
Db 783 KKSEA-----LESIKQKL-----TKQ---EEKKILKQDPETLSQETKIQ 819
QY 682 -KDIDNLIIFSLIKLYSLSDK-SLNWEPDSIASVISLIENTILGNKTSSEDEKTEKIR 739
Db 820 HEELNN-RIQTTVTTELQKVMKEALMTSLTVKDKLSKVSDSLKNSKSEFEKEN----- 873
QY 740 KLSYKYFFHYGIEFPDIQSGPDIVGNPPWEKTKFNETEFFSKHIPNYRKLGKIQEONII 799
Db 874 -----QKG-----KAAILDLEKTCELKH---QLQVQMENTL 902
QY 800 KQE-----ILSKD---NHPLSIBYNEKNSIIANNIYKDPDKFTSGGDPNLFYFVTF 851
Db 903 KEQKELKKSLEKEKASHQLKLELNSMOEQLIQAONTLKQNEK-----EEQQLQG 952
QY 852 NLKLIKEKGNLTIVPSAINWESSRILRKHIFARYKLVNIYOFENK----- 898
Db 953 NINELKQSSBQKKQKEALQGLKIAVLQK-----TELENKLOQLTQAAQELA 1001
QY 899 -----KRFKDVHSSFPKFAIFOLSNIKESTSSFKAKFMIOSSDNI----- 937
Db 1002 AEKEKISVLQNNVEKSGQETFKQLOSDPYGRESELLATROLKSVEEKLSLAQSDLSNRN 1061
QY 938 -----LKEITRDLKDSK-----DDAYKGIELNINQIKKLSPIQBSIIIEFKDNBEFTLINKM 988
Db 1062 QIGNQNKLIQELKATATLBQDSAKKEEQQLQ-ERCKALQDIQKE-----KSLKEKELVNEK 1116

QY 989 FSKSALGEGYIDFKK---GLDPSIKNRK-SLLKECNN-----KNLIF-----L 1028
Db 1117 -SKLAEIEIKCRQKEITKLNELKSHKLESIKETITNLKDAKQALLIQOKLELOGKADSL 1175
QY 1029 YSGANIHQPNSPFFEDKADAKESSKLLWIDKEDLEKVLTKDNQYQYTERVYFRAIASNTNR 1088
Db 1176 KAAVEQEKRNQOILKQVKEBEEL---KKEFIEKEAKLHSEIKKEVEGMKKHEENEAKL 1232
QY 1089 TWTISTLSPGNCYCVNSIYINDEKTPISLYKK 1119
Db 1233 TWTITALNENLGTVKKEWQSSQRRVSELEKQ 1263
RESULT 14
US-10-793-626-1660
; Sequence 1660, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1660
; LENGTH: 885
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-1660

Query Match 3.2%; Score 206.5; DB 6; Length 885;
Best Local Similarity 19.0%; Pred. No. 0.00075;
Matches 205; Conservative 188; Mismatches 313; Indels 371; Gaps 53;
QY 17 YKQLSKDFIKKENINKLKDFFILIKNKLFSIDNSTEANIESLLKYIFEBELNYSVBQOKA 76
Db 14 YKQLSKMEQSDVIVTVSDI-----DHYTEDN-----QRLDERLNLHLSQO-- 54
QY 77 QGIEGVESRVDILL-----FENDKDKASFNNKLKEAKNNEPIEDILLIAEVRKPT 129
Db 55 AEKEGQQAQINQLQKYGKGRQONDYDIEKLANVELVKATENEYQLS-GKLNVLLEERKQ 113
QY 130 FSFDAK--DKLASEDOLYRLNOKYKHGILSNGKVRWLYDKSVLYGKRYIEFNFXX 187
Db 114 SETNARYEBELNLESQDISIKNEKAQNEKLLA-----ELKNQKQKLNKEVQLESLLYI 168
QY 188 IEKEEYKEQWFLVIYILIRKERYLKTNSVIEVEKEQIAKEKEIITQTL--KEILYERP 245
Db 169 SDEQHDKELEE-----IKNSYITLMS-----EQSDVNDIRFLEHTINENEAKSRL 215
QY 246 DSQIV--FKIKNY-----DKFEKLSOK-----EITQHI--LASILEESIIFILRFFIA 292
Db 216 DSRLEAFNQLKDIQONTQTOKEYOSKSKMEKVEQNIQQLTQDTS----- 264
QY 293 YIEDNDIDFKKILQ--ENKLYRSSISFRYFFYDENTKKLEYKKIITIFNLDDKGSDAIKF 350
Db 265 -----KRLISEYENKLYQA-----YRYNEKLSRID-----SLATOEBEDYTYF 302
QY 351 PVFNGG--LFSDEK----- 362
Db 303 --FNGVKHILKAKOKELRGHGAEVAEINVPSMTQAIETALGASLQHVIVDNEKDGROA 360
QY 363 VKYLNNEGL-----LSISIEIEILVQMLFPEEKNIKD-----EKFKYSRLDPKSPGELY 412
Db 361 IQYLKQGRGLGRATFLPLNVIPQPHV-----AADINDVARGSQGFINIAS-DAINVSARY 413

413 ETLEVDLRADTTVRIRIIEEDGVYLRTEEBLEN--KVNKIATYKGNIIYTSRSLDRK 470
414 OMIEI---NLGNTI---IVEN---LKHANELARAIRYRTRIVT-LEGDVNVPFGSM--T 461
471 KSGAYTTPDLDTFMVISIEEQLTKSPLDIKIIDSNGSGHFLISCLDYLTEKVVYEL 530
462 GGGARKTKILSQKDELSTMRNLE-----DYQRQTAEFE- 496
531 DPEDVK---KELDEYRVIIIESEBY-DVQDSI-SKELVLRMLLKXCIYGVNDINPISV 585
497 RQFKEQKTAEOQLSEQY---FSASQOYNLKEQVHHHELEDRKLTQ----- 540
586 EITMLSWINTPIFGTPLSFIEHHIKAGNALLGYTKDEFPDIVKKKFSFGPSLPKKRIKE 645
541 -----EAHLKNEHEEFEEKNDGYQSDXS-----KETLKE 570
646 IITILEDIYQIKINGINTTKEDIEKSIYKYEESKDIDNLRIPSLIKLYSLSPDKSL 705
571 KQHLEIOQLKQL-----ESDIERYTQISKEGKAS-----THQTOQQL 610
706 NMEFSDIASVISLIEN-----ILGNKTSSEDEKIEKIRKLSVYKFFHYGIEFPDIOE 759
611 HQQSDLAIVKRIKQKVYERLDKQSDSEROKIEVNEKI-----KLFN----- 656
760 GPDIVIGNPWEKTK-----FNE--TEPFSKHIPNRYKLGIEKQNIUK--QEI 803
657 -SDMMGDAFEKLEQIOQOENVRQNLQOQLSDSEROKIEVNEKI-----KLFN----- 715
804 LSKDNHPLSIEVNEEK-----NSIIANNIYKFPKCTSGGDPNLFYFTFNKLJK 857
716 LSIENHYQIKAKQSKLDVLINHAIDHNDY-----PARYKANVIYQFENKGRKGVHSSP 908
858 EKNGLTYLVPISAIWNESSRILKHHI-----PARYKANVIYQFENKGRKGVHSSP 908
752 ERARMEYSDDETIDN-----LKKVKLTQMTIDELGPVNLNAIEQFE-----ELNERY 799
909 KFAIFOLSNIKESTSSFKAKFMIOSSDNILKEITRDLKSKDDAYKGIENLNIQIKK 965
800 TFLNEQRTDLREAK-----ETLEQIHEMDKEVGRFKTTFHVAQDHFTTVPK 847

RESULT 15
US/11/062
; Sequence 4, Application US/11062471A
; Publication No. US20050255093A1
; GENERAL INFORMATION:
; APPLICANT: SHONE, Clifford Charles
; APPLICANT: SUTTON, John Mark
; APPLICANT: HALLIS, Bassam
; APPLICANT: SILMAN, Nigel
; TITLE OF INVENTION: Delivery of Superoxide Dismutase to Neuronal Cells
; FILE REFERENCE: 1581.0800001
; CURRENT APPLICATION NUMBER: US/11/062,471A
; CURRENT FILING DATE: 2005-02-22
; PRIOR APPLICATION NUMBER: 09/831,050
; PRIOR FILING DATE: 1999-11-05
; PRIOR APPLICATION NUMBER: PCT/GB99/03699
; PRIOR FILING DATE: 1999-11-05
; PRIOR APPLICATION NUMBER: GB 9824282.9
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1070
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Construct comprising Mn-SOD from B. stearothermophilus, a linker,
US/11/062,471A-4

Query Match 3.2%; Score 206.5; DB 7; Length 1070;
Best Local Similarity 20.3%; Pred. No. 0.00094;
Matches 202; Conservative 171; Mismatches 353; Indels 271; Gaps 51;

Search completed: January 24, 2006, 20:55:47
Job time : 30.4381 secs

36 PFILINKKLFSDIDNSTEANTIESLLKIFYEELNYSVBOQKAGQIEGVSVDILLFENDK 95
234 FFIADKNS-FS-DDLSKNERIEYNTQSNYIENDFPINE-----LILDTD- 275
96 DKASFNKLEAKKNEPIEDILIIAIEVKRPTFSFAKDKLKESEDOLYRYLNOYQKH 155
276 ----LISKIELPSENTESLTFNVDPVYVEKQP-----AIKKI FTDENTIFQYL--YSQT 324
156 YGI-LSNGKVMRLYDKSKVLYGEXYIIFNFXXKIEBEYKQEWFLVFIYLRKERYLK 214
325 PFLDRIIDSLTSSPDDA-LLFSNKVYSPFSDMYIKTANKVVEAGLPAGVV-----KQIV 377
215 TSNVIEVEKEQIAKEKEIIQKTLKELIYERPDOSIVF-----KIAKNIYDKFK 263
378 NDFVTEANK-----SNTMDKI-----ADISLIVPYIGALNVGNSTAKGNFENAF 423
264 LSGKEITQHILASILEESIIFILRIPFI-AYIED-NDIPFKKILOENKL----- 309
424 IAGAS-----ILLEPIPELLIPVVGAPLLESYIDNKNKIITL--DNALTKRNEKWSMDYG 477
310 -----YRSSISPRYPFYDENTKKKLEY-----KKII-----TIFNLLDKGSDAIKPPVFN 355
478 LIVAQWLTNTQVFTYIKEGMYKALNYQAALIEEIIKYRYNIYSEKEKSNINIDFNDINS 537
356 GLFSEDKVKYLNNEGL-LSISEIEILVKQLFFBEKNIKDEKFKVYKSLRDPKSFGLYET 414
538 KL-----NEGNOAIDNNFI-----NGCSVSYLKKMKMPLAVBK 573
415 LLEVDLRADTTVHRIEDGVYLRTEEBLENKVNKIATYKGNIIYTSRSLDRKSGA 474
574 LUDFOITLKUNLIDENKLYIGS-ABEYKSKVNK-----YLK-----TIMFDL-----S 620
475 YTPDLDTFMVISIEEQLTKSPLDIKIIDSNGSGHFLISCLDYLTEKVMYLDKPE 534
621 IYTNDTIL-IEFNKYNSEILANNILNRYKDN-----LIDLSEY-----GAKVEYVD 668
535 DVKELDEBYRVIIIESEYVDQDSISKELVLKMLKXCIYGVNDINPISVEITMLSLMI 594
669 GV-ELNDKNQPKLTASSANSKIRVTQNNIIFNSVFLDFSV-----SPWI 711
595 NTFIFGTPLSFIEHHIKAGNALLGYTKDEFPDIVKKKFSFGPSLPKKRIKEIITILEDIY 654
712 RIPKY-----KNDGIQNTYHNEYTTINCMKNNSGWMKISIRGNRIITWLLD---- 756
655 QKIKGINDTFKEDIEKSKKIYKYESKDIDNL--RIIFSLI-----KLXSLSPDKSL 705
757 --ING-----XTKSVFFEYNIREDISEYINRWFFVTITNNLNNAKIY-INGKLES 803
706 NMEFSDIASVISLIE-----NILGNKTSSEDEKIEKIRKLSVY 745
804 NTDIKDIREVIANGELIPKLDGIDIDRTQFIWMKYFSIFNTELSQSNIEERYKIQSYEVL 863
746 KFFHVGIEPDIQEGFDIVGN-----PPWE---KTFENETEFESKHIPNRYKL 791
864 KDF-WGNPLMYNKYMFNAGNKSIIKLKDSPPVGEILTRSKYNQ-----SKYI-NYRDL 918
792 GIKEQNIKOEILSKDNHPLSIEYNEEKNSIIANNIYKFPKCTSGGDPNLFYFTVF 851
919 YIGEFILIR-----KSN-----SOSINDD---IVRKEDIYVLDFF-----NLNQEWRY 960
852 NLKLIKEKGNLTYPVPSAIWNESSRILKHHIFAR--YKLANIYQFENKGRKGVHSSPK 909
961 TYKYFKKEEKLFLAPISDSDEFYNTIQIKEYDEQPTYSQQLFK-----KDBESTDE 1013
910 FAIFOLSNIKES---TSSFKAKFMIOSSDNILKEITR 943
1014 IGLIGHIRPYESGIVFEYKDYFCI--SKWYLKEVKR 1048

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:21:03 ; Search time 14.5258 Seconds
(without alignment)
1966.133 Million cell updates/sec

Title: US-10-688-058-18
Perfect score: 329
Sequence: 1 LAATPLVDVTGKLRITAKI.....GAIDPQTYHYKNEAIIEL 65

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*
9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	329	100.0	65	8	ADU98759 Borrelia
2	329	100.0	160	8	ADU98803 Borrelia
3	329	100.0	370	8	ADU98805 Borrelia
4	329	100.0	370	8	ADU98761 Borrelia
5	73	22.2	233	5	AB55144 Lactococc
6	67	20.4	188	8	ADT59925 Plant pol
7	65	19.8	284	6	ABM73458 Staphyloc
8	64.5	19.6	348	7	ABM90374 Rice abio
9	64.5	19.6	601	6	ABU33885 Protein e
10	64	19.5	138	8	ADY05830 Plant ful
11	64	19.5	148	8	ADT58407 Plant pol
12	64	19.5	193	8	ADY04902 Plant ful
13	64	19.5	193	8	ADY78687 Plant ful
14	64	19.5	193	8	ADY22628 Plant ful
15	64	19.5	193	8	ADY94536 Plant ful
16	63.5	19.3	301	8	ADN17923 Bacterial
17	63	19.1	380	6	ABU29425 Protein e
18	63	19.1	381	7	ADH87344 Enterococ
19	63	19.1	491	2	AAW09771 Pichia ac
20	62.5	19.0	402	8	ADN29219 Bacterial
21	62.5	19.0	404	7	ABM86837 Rice abio
22	62.5	19.0	445	9	ADY71564 Laccase p
23	62.5	19.0	1017	8	ADN41736 Bacterial
24	62	18.8	351	9	ABM93745 M. xanthu

25	62	18.8	580	6	AAO19831.	Aac19831 Bordetell
26	62	18.8	750	6	AAE36322	Aac36322 Arabidops
27	61.5	18.7	122	6	ABU24382	Abu24382 Protein e
28	61.5	18.7	302	4	AAG82943	Aag82943 S. epider
29	61.5	18.7	302	6	ABU43297	Abu43297 Protein e
30	61.5	18.7	327	7	ABO64173	AbO64173 Klebsiell
31	61.5	18.7	350	5	ABP38538	Abp38538 Staphyloc
32	61.5	18.7	350	8	ADS06181	AdS06181 Staphyloc
33	61.5	18.7	695	4	ABG28835	Abg28835 Novel hum
34	61	18.5	355	4	AAU35146	Aau35146 Enterococ
35	61	18.5	355	6	ABU29086	Abu29086 Protein e
36	61	18.5	355	8	ADH97192	Adh97192 E. faecal
37	61	18.5	355	8	ADH97194	Adh97194 E. faecal
38	61	18.5	355	9	AEC13547	Aec13547 Enterococ
39	61	18.5	355	9	AEC13549	Aec13549 Enterococ
40	61	18.5	371	7	ADH87288	Adh87288 Enterococ
41	61	18.5	384	4	AAG92653	Aag92653 C glutami
42	61	18.5	388	4	AAAB79449	Aab79449 Corynebac
43	61	18.5	388	4	AAAB79450	Aab79450 Corynebac
44	60.5	18.4	645	7	ABO76507	AbO76507 Pseudomon
45	60	18.2	144	3	AAAB58826	Aab58826 Breast an

ALIGNMENTS

RESULT 1
ADU98759
ID ADU98759 standard; protein; 65 AA.
XX
AC ADU98759;
XX
DT 24-FEB-2005 (first entry)
XX
DE Borrelia burgdorferi antigenic polypeptide seqid 18.
XX
KW antibacterial; vaccine; immune stimulation; immunity; antigen;
KW DNA library.
XX
OS Borrelia burgdorferi.
XX
FN WO2004103269-A2.
XX
PD 02-DEC-2004.
XX
PF 17-OCT-2003; 2003WO-US033056.
XX
PR 18-OCT-2002; 2002US-0419401P.
XX
PA (MACR-) MACROGENICS INC.
XX (TEXA) UNIV TEXAS SYSTEM.
XX
PI Sykes KF, Hale KS, Johnston SA;
XX
DR WPI; 2004-834155/82.
XX N-PSDB; ADU98758.
XX
PT Immunizing a subject against Borrelia burgdorferi infection comprises
XX providing to the subject at least one Borrelia antigen or its fragment.
XX
PS Claim 27; SEQ ID NO 18; 121pp; English.
XX
CC The invention describes a method of immunizing a subject comprising
CC providing to the subject at least one Borrelia antigen or its fragment to
CC induce an immune response. Also described are: an isolated polynucleotide
CC comprising a sequence having at least 17 contiguous nucleotides in common
CC with a sequence not given in the specification; an isolated polypeptide
CC having at least 5 consecutive amino acids of the sequence not given in
CC the specification; a vaccine composition comprising at least one Borrelia
CC antigen or at least one polynucleotide encoding a Borrelia antigen;
CC screening for at least one test polypeptide or test polynucleotide
CC encoding a polypeptide for an ability to produce an immune response;
CC preparing a vaccine; vaccinating a subject; treating a subject infected

XX DT 02-JUN-2005 (first entry)
 XX DE Rice abiotic stress responsive polypeptide SEQ ID NO:9096.
 XX KW abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.
 XX OS *Oryza sativa*.
 XX PN WO2003008540-A2.
 XX PD 30-JAN-2003.
 XX PF 21-JUN-2002; 2002WO-US019668.
 XX PR 22-JUN-2001; 2001US-0300112P.
 XX PR 24-AUG-2001; 2001US-0314662P.
 XX PR 26-SEP-2001; 2001US-0325277P.
 XX PR 21-NOV-2001; 2001US-0332132P.
 XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX PI Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;
 XX PI Moughamer T, Provart N, Ricke D, Zhu T;
 XX WPI; 2003-248011/24.
 XX New stress-responsive nucleic acid, useful for altering the
 PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
 PT stress, salt stress or osmotic stress.
 XX Claim 1; SEQ ID NO 9096; 89pp; English.
 CC The invention relates to novel abiotic stress responsive polynucleotides
 CC and polypeptides. Also disclosed are vectors, expression cassettes, host
 CC cells, and plants containing such polynucleotides. Also disclosed are
 CC methods for using the polynucleotides and polypeptides to alter the
 CC responsiveness of a plant to abiotic stress. The invention is useful in
 CC agriculture. The nucleic acid is useful for determining whether a test
 CC plant has been exposed to an abiotic stress condition. It is also useful
 CC for selecting an agent that alters abiotic stress regulated
 CC polynucleotide expression in a plant cell, and to identify a homolog or
 CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
 CC molecule and the polypeptide encoded by it are useful in altering the
 CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
 CC stress, osmotic stress or any of their combinations. The present sequence
 CC is used in the exemplification of the invention
 XX Sequence 348 AA;
 SQ
 Query Match 19.6%; Score 64.5; DB 7; Length 348;
 Best Local Similarity 33.9%; Pred. No. 12;
 Matches 20; Conservative 11; Mismatches 23; Indels 5; Gaps 3;
 QY 12 IGLRTA--KINFYSLNET--GLDGVPAFKEG-VDLGAGGAIDQFTYIKNEAIIEL 65
 DB 231 VGDLRVPRFKVSFYSEMEVNLKMGIGAFDVGKVDLSGMXGELVWVEKVMHRAVVEV 289
 RESULT 9
 ABU33885
 ID ABU33885 standard; protein; 601 AA.
 XX AC ABU33885;
 DT 19-JUN-2003 (first entry)
 XX Protein encoded by Prokaryotic essential gene #19412.
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KW Mycobacterium avium.
 XX OS
 XX

PN WO200277183-A2.
 XX PD 03-OCT-2002.
 XX PF 21-MAR-2002; 2002WO-US009107.
 XX PR 21-MAR-2001; 2001US-00815242.
 XX PR 06-SEP-2001; 2001US-00948993.
 XX PR 25-OCT-2001; 2001US-0342923P.
 XX PR 08-FEB-2002; 2002US-00072851.
 XX PR 06-MAR-2002; 2002US-0362699P.
 XX PA (ELIT-) ELITRA PHARM INC.
 XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 XX N-PSDB; ACA37755.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 61809; 1766pp; English.
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 601 AA;
 SQ
 Query Match 19.6%; Score 64.5; DB 6; Length 601;
 Best Local Similarity 32.7%; Pred. No. 23;
 Matches 17; Conservative 8; Mismatches 24; Indels 3; Gaps 1;
 QY 6 LVDDTVIGKLTAKIN---FYSLLNETGLDGVPAFKEGVDLGAGGAIDQFTY 54
 DB 257 LVADVPVCGLLSGVDSSLIIVGLLAARQGHLSIFSGFESAGGVKGDFQY 308
 RESULT 10
 ADY05830
 ID ADY05830 standard; protein; 138 AA.
 XX

CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040216190.

XX SQ Sequence 148 AA;

Query Match 19.5%; Score 64; DB 8; Length 148;
Best Local Similarity 35.0%; Pred. No. 4.7;
Matches 21; Conservative 8; Mismatches 25; Indels 6; Gaps 2;

QY 6 LVDDTVIGKLTAKINFYSLNETGLDG--VPAPKEGVLDAG---GAIDQFTVHYIKN 59
Db 57 VVDSSTDLRLVTAKBEFFHAILEDELKGAIVLVYANKQDLPGLDAAITSELEHLKIKS 116

RESULT 12

ID ADY04902
ADY04902 standard; protein; 193 AA.

XX AC ADY04902;

XX DT 21-APR-2005 (first entry)

XX DE Plant full length insert polypeptide seqid 60717.

XX KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
XX protein content.

XX OS Unidentified.

XX OS US2004034888-A1.

XX PD 19-FEB-2004.

XX PF 28-APR-2003; 2003US-00425114.

XX PR 06-MAY-1999; 99US-00304517.

XX PR 05-NOV-2001; 2001US-00985678.

XX PA (LIUJ/) LIU J.

XX PA (ZHOU/) ZHOU Y.

XX PA (KOVA/) KOVALIC D K.

XX PA (SCRE/) SCREEN S E.

XX PA (TABR/) TABASKA J E.

XX PA (CAOY/) CAO Y.

XX PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;

XX DR WPI; 2004-180133/17.

XX PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.

XX PS Claim 1; SEQ ID NO 60717; 15pp; English.

XX CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring

CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
XX invention.

XX SQ Sequence 193 AA;

Query Match 19.5%; Score 64; DB 8; Length 193;
Best Local Similarity 35.0%; Pred. No. 6.6;
Matches 21; Conservative 8; Mismatches 25; Indels 6; Gaps 2;

QY 6 LVDDTVIGKLTAKINFYSLNETGLDG--VPAPKEGVLDAG---GAIDQFTVHYIKN 59
Db 102 VVDSSTDLRLVTAKBEFFHAILEDELKGAIVLVYANKQDLPGLDAAITSELEHLKIKS 161

RESULT 13

ID ADX78687

ADX78687 standard; protein; 193 AA.

XX AC ADX78687;

XX DT 21-APR-2005 (first entry)

XX DE Plant full length insert polypeptide seqid 48053.

XX KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
XX protein content.

XX OS Unidentified.

XX OS US2004034888-A1.

XX PD 19-FEB-2004.

XX PF 28-APR-2003; 2003US-00425114.

XX PR 06-MAY-1999; 99US-00304517.

XX PR 05-NOV-2001; 2001US-00985678.

XX PA (LIUJ/) LIU J.

XX PA (ZHOU/) ZHOU Y.

XX PA (KOVA/) KOVALIC D K.

XX PA (SCRE/) SCREEN S E.

XX PA (TABR/) TABASKA J E.

XX PA (CAOY/) CAO Y.

XX PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;

XX DR WPI; 2004-180133/17.

XX PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.

XX PS Claim 1; SEQ ID NO 48053; 15pp; English.

XX CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide

[illegible]

The invention describes a recombinant DNA construct comprising a polynucleotide consisting of a sequence encoding an amino acid sequence available in electronic form from the US patent office at [ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888](http://seqdata.uspto.gov/sequence.html?DocID:2004034888). The polynucleotide of the invention are also useful in physical arrays of molecules and as plant breeding markers. The recombinant DNA construct is useful for improving plant tolerance to cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or pests, for manipulating growth rate in plant cells by modification of the cell cycle pathway, for conferring increased resistance to plant disease, for producing galactomannan, lignin or plant growth regulators, for increasing the rate of homologous recombination in plants, for improving yield by modification of photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake or by providing improved plant growth and development under at least one stress condition or for modifying seed oil or protein yield and/or content. This is the amino acid sequence of a plant full length insert polypeptide that can be used in the recombinant DNA construct of the invention.

Sequence 193 AA;

Query Match 19.5%; Score 64; DB 8; Length 193;
Best Local Similarity 35.0%; Pred.No. 6.6;
Matches 21; Conservative 8; Mismatches 25; Indels 6; Gaps 2;

Qy 6 LVDDTVIGKLTAKINFYSLNETGLDG--VPAPKEGVDLG---GAIDEQFTYHYIKN 59
Db 102 VVDSSTDRLVTAKEFHAILEDEELXGAVVLVYANKQDLPALDDAAITESLELHKKS 161
::: : | ||| | ::|| : : : : | : | : | : |
: : : : : : : : : : : : : : : :

RESULT 15
ADX94536

ID ADX94536 standard; protein; 193 AA.
AC ADX94536;
XX
DT 21-APR-2005 (first entry)
DE Plant full length insert polypeptide seqid 57200.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
protein content.
XX
OS Unidentified.
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PP 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVALIC D K.) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA ('TABAJ') TABASKA J E.
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX
PT New recombinant DNA construct, useful for improving plant tolerance to

This Page Blank (uspto)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	207	62.9	356	2	F70212	hypothetical prote
2	73	22.2	233	2	B86849	regulatory protein
3	67	20.4	337	2	B84335	hypothetical prote
4	64.5	19.6	307	2	AE0962	D-serine deaminase
5	64	19.5	284	2	F90016	conserved hypothet
6	63.5	19.3	301	2	F75419	succinyl-CoA synth
7	63.5	19.3	345	2	C84489	probable replicati
8	63.5	19.3	709	2	E90210	hypothetical prote
9	63	19.1	491	2	S42678	killer toxin - Pic
10	62	18.8	703	2	D81365	carbon starvation
11	62	18.8	781	2	T02272	hypothetical prote
12	61	18.5	785	2	E69884	hypothetical prote
13	60.5	18.4	612	2	S61116	probable membrane
14	60	18.2	233	2	A64307	probable transcrip
15	60	18.2	301	2	S52142	glycosyl transferas
16	60	18.2	375	2	B87394	hypothetical prote
17	60	18.2	397	2	B69378	probable acyl-CoA
18	60	18.2	540	2	B83850	D-3-phosphoglycer
19	59	17.9	348	2	AD2247	hypothetical prote
20	59	17.9	402	2	D69530	probable acyl-CoA
21	59	17.9	1228	2	C98219	proline dehydrogen
22	59	17.9	1228	2	AG3067	proline dehydrogen
23	59	17.9	1978	2	S77257	hypothetical prote
24	58.5	17.8	218	2	S30287	regulatory protein
25	58.5	17.8	218	2	S07359	regulatory protein
26	58.5	17.8	309	2	AH3122	IS3 family transpo
27	58.5	17.8	309	2	G98164	IS3 family orf1B pr
28	58.5	17.8	311	2	A65010	D-serine deaminase
29	58	17.6	182	2	S41938	ADP-ribosylation f

A;Cross-references: UNIPROT:Q9CEP1;

A;Cross-references: UNIPROT:Q9CEP1; UNIPARC:UPI00000C6B13; GB:AE005176; PID:gi2724820; P
A;Experimental source: strain IL1403

Query Match 19.6%; Score 64.5; DB 2; Length 307;
Best Local Similarity 39.5%; Pred. No. 7.8;
Matches 17; Conservative 7; Mismatches 16; Indels 3; Gaps 1;

RESULT 7
C84489

probable replication protein A1 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: C84489
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.;
Euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: C84489
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-345 <STO>
A:Cross-references: UNIPROT:Q9SK24; UNIPARC:UPI000009F829; GB:AE002093; NID:G4558677; P1
C:Genetics:
A:Gene: At2G09960
A:Map position: 2

Query Match 19.3%; Score 63.5; DB 2; Length 345;
Best Local Similarity 26.4%; Pred. No. 12;
Matches 19; Conservative 10; Mismatches 20; Indels 23; Gaps 2;

Qy 2 AATPLVDDTVIGKLTAKINFY-----SLLNETGLDGVPAKGVLDL 43
Db 103 ACQQLTGDGVCLIRFAKINLYNDRSVNSFVSQVPVDSLTDLGL-----FKQSIP 157
Qy 44 AGGAIDEQFTYH 55
Db 158 DGLTLESSGSF 169

RESULT 8
E90210
hypothetical protein purL [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: E90210
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.
arrett, R.A.; Ragan, M.A.; Sensesen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A39139
A:Accession: E90210
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-709 <KUR>
A:Cross-references: UNIPROT:Q9UX24; UNIPARC:UPI0000132B43; GB:AE006641; NID:G13813798; F
C:Genetics:
A:Gene: purL
C:Superfamily: phosphoribosylformylglycinamide synthase component II

Query Match 19.3%; Score 63.5; DB 2; Length 709;
Best Local Similarity 46.5%; Pred. No. 26;
Matches 20; Conservative 5; Mismatches 17; Indels 1; Gaps 1;

Qy 5 PLVDDTVIGKLTAKINFYSLNETGLDGVPAKGVLDL 47
Db 168 PLVDDTVIGKLTAKIK-PSIVDKAGLVLAAGLVGVDLGG 209

RESULT 9
S42678
killer toxin - Pichia acaciae plasmid pPacl-2
C:Species: Pichia acaciae
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S42678
R:Boelen, P.L.; Eastman, E.M.; Cihak, P.L.; Hayman, G.T.
Yeast 10, 403-414, 1994
A:Title: Isolation and sequence analysis of a gene from the linear DNA plasmid pPacl-2
A:Reference number: S42678; MUID:94287718; PMID:8017110
A:Accession: S42678

A:Molecule type: DNA
A:Residues: 1-491 <BOL>
A:Cross-references: UNIPROT:Q01498; UNIPARC:UPI000006B991; GB:U02596; NID:G435925; PIDN:
A:Experimental source: NRRL Y-18665
C:Genetics:
A:Genome: plasmid pPacl-2
C:Keywords: glycoprotein
F:1135,205,241,263,291,354,387/Binding site: carbohydrate (Asn) (covalent) #status predic
Query Match 19.1%; Score 63; DB 2; Length 491;
Best Local Similarity 25.4%; Pred. No. 20;
Matches 16; Conservative 11; Mismatches 16; Indels 20; Gaps 2;

Qy 9 DTVIG-----KLRTAKINFYSLNETGLDGVPAKGVLDL 62
Db 278 DNIYGYMQLQRIINNTENFSLTNEVTXGVP-----DHFYFYAKNDPM 323
Qy 63 IEL 65
Db 324 IQI 326

RESULT 10
D81365
carbon starvation protein A homolog Cj0917c [similarity] - Campylobacter jejuni (strain
C:Species: Campylobacter jejuni
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: D81365
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillin
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrel
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: D81365
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-703 <PR>
A:Cross-references: UNIPROT:Q9PPI6; UNIPARC:UPI000000CID90; GB:AL139076; GB:AL111168; NID
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: cstA; Cj0917c
C:Superfamily: carbon starvation protein

Query Match 18.8%; Score 62; DB 2; Length 703;
Best Local Similarity 38.2%; Pred. No. 38;
Matches 21; Conservative 6; Mismatches 20; Indels 8; Gaps 3;

Qy 17 TAKINFYSLNETGLDGVPAKGV-----DLAGAIDEQFTYHY-INKNEAIIEL 65
Db 428 TTNIGYTIILSTG--GAPTFAIGVALTLHELFGVLDLMAFWYHFAILFEALFIL 480

RESULT 11
T02272
hypothetical protein T1308.9 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004
C:Accession: T02272
R:Vysotskaia, V.S.; Schwartz, J.R.; Kwan, A.; Toriumi, M.; Yu, G.; Oji, O.; Liu, S.; Li,
rtz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel, N.
submitted to the EMBL Data Library, June 1998
A:Description: Arabidopsis thaliana chromosome 1 BAC T13D8 sequence.
A:Reference number: 214649
A:Accession: T02272
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-781 <VVS>
A:Cross-references: UNIPROT:O80743; UNIPARC:UPI000000A9B72; EMBL:AC004473; NID:G3108025;
C:Genetics:
A:Gene: ATSP:T13D8.9
A:Map position: 1
A:Introns: 75/3; 126/3; 259/1; 478/1; 673/3

Query Match 18.8%; Score 62; DB 2; Length 781;
 Best Local Similarity 48.1%; Pred. No. 43;
 Matches 13; Conservative 4; Mismatches 8; Indels 2; Gaps 1;

QY 27 NETGLDGVPAFKEGVLDLAGAI--DEQ 51
 DB 177 NETGKGESKIKENIDIANSAVLTDDE 203

RESULT 12
 E69884
 Hypothetical protein ymcC - Bacillus subtilis
 C:Species: Bacillus subtilis
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C:Accession: E69884
 R:Kunst, P.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertez
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
 iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.P.
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
 A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: E69884
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-185 <KUN>
 A:Cross-references: UNIPROT:O31780; UNIPARC:UPI000006042F; GB:Z99112; GB:AL009126; NID:9
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: ymcC
 C:Superfamily: Bacillus subtilis hypothetical protein ymcC

Query Match 18.5%; Score 61; DB 2; Length 185;
 Best Local Similarity 28.6%; Pred. No. 11;
 Matches 18; Conservative 11; Mismatches 26; Indels 8; Gaps 2;

QY 1 LAATPLVDVTGKLTAKINFYSLNLTGLDGVPAFKEGVLDAGG-----IDQFTYH 55
 DB 39 LALTFVID---LILLAATGVLDYRGASATAAHGIAAVVIGISYAYGKQIQWADEKFOY 95

QY 56 YIK 58
 DB 96 VTK 98

RESULT 13
 S61116
 Probable membrane protein YPL070w - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein LPF14w
 C:Species: Saccharomyces cerevisiae
 C>Date: 23-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
 C:Accession: S61116
 R:Hall, J.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Storms, R.K.; Vo, D.H.; Wa
 submitted to the EMBL Data Library, August 1995
 A:Description: The sequence of Saccharomyces cerevisiae chromosome XVI left arm.
 A:Reference number: S59677
 A:Accession: S61116
 A:Molecule type: DNA
 A:Residues: 1-612 <HAL>
 A:Cross-references: UNIPROT:Q02866; UNIPARC:UPI000006A497; EMBL:U41849; NID:g1147608; PI
 C:Genetics:
 A:Gene: MIPS:YPL070w
 A:Cross-references: SGD:S0005991
 A:Map position: 16L

C:Superfamily: Saccharomyces cerevisiae probable membrane protein YPL070w
 C:Keywords: transmembrane protein
 F:356-372/Domain: transmembrane #status predicted <TM>

Query Match 18.4%; Score 60.5; DB 2; Length 612;
 Best Local Similarity 35.0%; Pred. No. 49;
 Matches 21; Conservative 10; Mismatches 18; Indels 11; Gaps 4;

QY 12 IGKLRTA-KINFYSLN-ETGLDGVPAFKEGVLDLAGAIDQFTYHYI---KNEAIIEL 65
 DB 286 LSKLSTWNKLNLIKFLSLNNGIDPEKFEIKDIL-----YEFTYHSISPEKIKALKKL 340

RESULT 14
 A64907
 Probable transcription regulator yneJ - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 05-Oct-2004
 C:Accession: A64907; S41477; S35950
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A>Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: A64907
 A>Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-293 <BLAT>
 A:Cross-references: UNIPROT:P77309; UNIPARC:UPI000013BB1C; GB:AE000250; GB:U00096; NID:9
 A:Experimental source: strain K-12, substrain MG1655
 R:Cecchini, E.; Dominy, P.J.; Geri, C.; Kaiser, K.; Sentry, J.; Milner, J.J.
 Nucleic Acids Res. 21, 5742-5747, 1993
 A>Title: Identification of genes up-regulated in dedifferentiating Nicotiana glauca pit
 A:Reference number: S41477; MUID:94111998; PMID:8284223
 A:Accession: S41477
 A>Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 'MP', 173-186; 190-274 <CEC>
 A:Cross-references: UNIPARC:UPI0000178EDE; UNIPARC:UPI0000178EDF; EMBL:X74318; NID:94176
 A:Notice: the authors state that the gene has been cloned from Nicotiana glauca, most like
 C:Genetics:
 A:Gene: yneJ
 C:Superfamily: HTH-type transcriptional regulator, LysR family, MetR type
 C:Keywords: DNA binding; transcription regulation

Query Match 18.2%; Score 60; DB 2; Length 293;
 Best Local Similarity 35.3%; Pred. No. 25;
 Matches 12; Conservative 9; Mismatches 11; Indels 2; Gaps 1;

QY 6 LVDDTVIGKLTAKINFYSLNLTGLDGVPAFKE 39
 DB 130 MLDGVLEKLNAAFID--GPINHTAIDGIPVRE 161

RESULT 15
 S52142
 Glycosyl transferase amsB - Erwinia amylovora
 C:Species: Erwinia amylovora
 C:Date: 15-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
 C:Accession: S61895; S52142
 R:Bugert, P.; Geider, K.
 Mol. Microbiol. 15, 917-933, 1995
 A>Title: Molecular analysis of the ams operon required for exopolysaccharide synthesis
 A:Reference number: S61891; MUID:95319333; PMID:7596293
 A:Accession: S61895
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-301 <BUZ>
 A:Cross-references: UNIPROT:Q46632; UNIPARC:UPI0000125A31; EMBL:X77921; NID:g600426; PI
 C:Superfamily: Neisseria meningitidis glycosyl transferase A

Query Match 18.2%; Score 60; DB 2; Length 301;
 Best Local Similarity 29.3%; Pred. No. 25;

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:27:53 ; Search time 13.9671 Seconds
(without alignments)
3283.383 Million cell updates/sec

Title: US-10-688-058-18
Perfect score: 329
Sequence: 1 LAATPLVDDTVIGKLTAKI.....GAIDEQTYHYKNEAIIEL 65

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	329	100.0	370	Q5XYW3	BORR
2	329	100.0	370	Q9R2W7	BORR
3	329	100.0	370	Q9S096	BORR
4	329	100.0	370	Q9S0C8	BORR
5	329	100.0	370	Q9S0F9	BORR
6	316	96.0	362	Q9S051	BORR
7	304	92.4	370	Q9S003	BORR
8	298	90.6	373	Q5XZ68	BORR
9	207	62.9	356	Q5O939	BORR
10	126	38.3	156	Q6AS11	BORR
11	81	24.6	250	Q54MG3	DICDI
12	73.5	22.3	402	Q97BQ1	THEVO
13	73	22.2	233	MECA	LACIA
14	69.5	21.1	1046	Q8EBM6	SHEON
15	68.5	20.8	6858	Q7QUM1	ITALA
16	67	20.4	337	Q9HP24	HALSA
17	67	20.4	394	Q6L1Q8	PICTO
18	67	20.4	745	Q4X0S2	ASFPF
19	66	20.1	1592	Q5AZJ8	EMENI
20	65.5	19.9	174	Q59PV8	CANAL
21	65	19.8	284	Q5HE01	STAAC
22	65	19.8	410	Q8DV08	STRMU
23	65	19.8	445	Q90681	ASTPE
24	65	19.8	703	Q4HDM0	CAMCO
25	65	19.8	1315	Q9R9T7	PSEP
26	65	19.8	1317	Q88D80	PSEP
27	64.5	19.6	307	Q5PKR5	SALPA
28	64.5	19.6	307	Q82L21	SALTI
29	64.5	19.6	307	Q8ZL10	SALTI
30	64.5	19.6	312	Q57134	SALCH
31	64.5	19.6	601	Q740D0	MYCPA

32 64.5 19.6 1517 2 Q4HLT3 CAMLA
33 64 19.5 182 2 Q6YW27 ORISA
34 64 19.5 284 2 Q6GEM6 STAAR
35 64 19.5 284 2 Q7A485 STAAAN
36 64 19.5 284 2 Q9S666 STAAAM
37 64 19.5 589 2 Q4T1A4 AZOVI
38 63.5 19.3 248 2 Q5JXW0 HUMAN
39 63.5 19.3 301 2 Q9RUY2 DEIRA
40 63.5 19.3 345 2 Q9SK24 ARATH
41 63.5 19.3 709 1 PURL SULSO
42 63 19.1 138 2 Q73M04 TREDE
43 63 19.1 377 2 Q835P7 ENTEFA
44 63 19.1 424 2 Q53KS8 ORISA
45 63 19.1 451 2 Q7XMI6 ORISA

ALIGNMENTS

RESULT 1
Q5XYW3 BORGA PRELIMINARY; PRT; 370 AA.
AC Q5XYW3,
DT 25-OCT-2004 (Tremblrel. 28, Created)
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
DE Hypothetical protein.
GN ORFNames=BGPI68;
OS Borrelia garinii PBi.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=290434;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBi;
RA Glockner G., Lehmann R., Romualdi A., Pradella S.,
RA Schulte-Spechtel U., Schilhabel M., Wilke B., Suhnel J., Platzner M.;
RT "Comparative analysis of the Borrelia garinii genome.";
RL Nucleic Acids Res. 32:6038-6046(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBi;
RA Gloeckner G., Schilhabel M., Lehmann R., Platzner M.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY722923; AAU86019.1; -; Genomic_DNA.
DR InterPro; IPR008505; DUF787.
DR Pfam; PF05619; DUF787; 1.
KW Hypothetical protein.
SQ SEQUENCE 370 AA; 42173 MW; 62FB9015CEBFC9C3 CRC64;

Query Match 100.0%; Score 329; DB 2; Length 370;
Best Local Similarity 100.0%; Pred. No. 6.6e-30;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LAATPLVDDTVIGKLTAKINFYSLNETGLDGVPAFKGVLDLGGAGIDEQTYHYKNE 60
DB 211 LAATPLVDDTVIGKLTAKINFYSLNETGLDGVPAFKGVLDLGGAGIDEQTYHYKNE 270

QY 61 AIIEL 65

DB 271 AIIEL 275

RESULT 2

Q9R2W7 BORBU PRELIMINARY; PRT; 370 AA.
AC Q9R2W7;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-FEB-2005 (Tremblrel. 29, Last annotation update)
DE Hypothetical protein BB11 (Hypothetical protein BBP11) (Hypothetical protein BBS11).
GN OrderedLocusNames=BB11, BBP11, BBS11;

OS Borrelia burgdorferi (Lyme disease spirochete).
 OG Plasmid cp32-8, plasmid cp32-1, and plasmid cp32-3.
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
 OC Borrelia burgdorferi group.
 ON NCBI_TaxID=139;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 35210 / B31; PLASMID=cp32-8, cp32-1, and cp32-3;
 RX MEDLINE=20138354; PubMed=10672174;
 RA Casjens S., Palmer N., van Vugt R., Huang W.M., Stevenson B., Rosa P.,
 RA Lathigra R., Sutton G.G., Peterson J.D., Dodson R.J., Haft D.H.,
 RA Hickey E.K., Gwinn M.L., White O., Fraser C.M.;
 RA "A bacterial genome in flux: the twelve linear and nine circular
 RT extrachromosomal DNAs in an infectious isolate of the Lyme disease
 RT spirochete Borrelia burgdorferi";
 RL Mol. Microbiol. 35:490-516(2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 35210 / B31; PLASMID=cp32-8, cp32-1, and cp32-3;
 RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
 RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
 RA Richardson D.L., Peterson J.D., Kevlavage A.R., Quackenbush J.,
 RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
 RA Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A.,
 RA Artiach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
 RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
 RA "Genomic sequence of a Lyme disease spirochete, Borrelia
 RT burgdorferi";
 RL Nature 390:580-586(1997).
 DR EMBL; AE001575; AAF07434.1; -; Genomic DNA.
 DR EMBL; AE001576; AAF07479.1; -; Genomic DNA.
 DR EMBL; AE001580; AAF07651.1; -; Genomic DNA.
 DR TIGR; BB11; -;
 DR TIGR; BBP1; -;
 DR TIGR; BBS1; -;
 DR InterPro; IPR008505; DUF787.
 DR Pfam; PF05619; DUF787; 1.
 KW Complete proteome; Hypothetical protein; Plasmid.
 SQ SEQUENCE 370 AA; 42102 MW; D071E1723FA9767C CRC64;
 Query Match 100.0%; Score 329; DB 2; Length 370;
 Best Local Similarity 100.0%; Pred. No. 6.6e-30;
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LAATPLVDDTVIGKLTAKINFYSLNETGLDGPAPKGVLAGGAIDEQFTYHYKNE 60
 Db 211 LAATPLVDDTVIGKLTAKINFYSLNETGLDGPAPKGVLAGGAIDEQFTYHYKNE 270
 QY 61 AIIEL 65
 Db 271 AIIEL 275
 RESULT 3
 Q9S096 BORBU PRELIMINARY; PRT; 370 AA.
 AC Q9S096;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein BB011.
 GN OrderedLocustNames=BB011;
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OG Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
 OC Borrelia burgdorferi group.
 ON NCBI_TaxID=139;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 35210 / B31;
 RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
 RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
 RA Richardson D.L., Peterson J.D., Kevlavage A.R., Quackenbush J.,
 RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
 RA Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A.,
 RA Artiach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
 RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
 RA "Genomic sequence of a Lyme disease spirochete, Borrelia
 RT burgdorferi";
 RL Nature 390:580-586(1997).
 DR EMBL; AE001575; AAF07434.1; -; Genomic DNA.
 DR EMBL; AE001576; AAF07479.1; -; Genomic DNA.
 DR EMBL; AE001580; AAF07651.1; -; Genomic DNA.
 DR TIGR; BB11; -;
 DR TIGR; BBP1; -;
 DR TIGR; BBS1; -;
 DR InterPro; IPR008505; DUF787.
 DR Pfam; PF05619; DUF787; 1.
 KW Complete proteome; Hypothetical protein; Plasmid.
 SQ SEQUENCE 370 AA; 42102 MW; D071E1723FA9767C CRC64;
 Query Match 100.0%; Score 329; DB 2; Length 370;
 Best Local Similarity 100.0%; Pred. No. 6.6e-30;
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LAATPLVDDTVIGKLTAKINFYSLNETGLDGPAPKGVLAGGAIDEQFTYHYKNE 60
 Db 211 LAATPLVDDTVIGKLTAKINFYSLNETGLDGPAPKGVLAGGAIDEQFTYHYKNE 270
 QY 61 AIIEL 65
 Db 271 AIIEL 275
 RESULT 4
 Q9S0C8 BORBU PRELIMINARY; PRT; 370 AA.
 AC Q9S0C8;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein BBM11.
 GN OrderedLocustNames=BBM11;
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OG Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
 OC Borrelia burgdorferi group.
 ON NCBI_TaxID=139;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 35210 / B31;
 RX MEDLINE=20138354; PubMed=10672174;
 RA Casjens S., Palmer N., van Vugt R., Huang W.M., Stevenson B., Rosa P.,
 RA Lathigra R., Sutton G.G., Peterson J.D., Dodson R.J., Haft D.H.,
 RA Hickey E.K., Gwinn M.L., White O., Fraser C.M.;
 RA "A bacterial genome in flux: the twelve linear and nine circular
 RT extrachromosomal DNAs in an infectious isolate of the Lyme disease
 RT spirochete Borrelia burgdorferi";
 RL Mol. Microbiol. 35:490-516(2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 35210 / B31;
 RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,

RA Casjens S., Palmer N., van Vugt R., Huang W.M., Stevenson B., Rosa P.,
 RA Lathigra R., Sutton G.G., Peterson J.D., Dodson R.J., Haft D.H.,
 RA Hickey E.K., Gwinn M.L., White O., Fraser C.M.;
 RA "A bacterial genome in flux: the twelve linear and nine circular
 RT extrachromosomal DNAs in an infectious isolate of the Lyme disease
 RT spirochete Borrelia burgdorferi";
 RL Mol. Microbiol. 35:490-516(2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 35210 / B31;
 RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
 RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
 RA Richardson D.L., Peterson J.D., Kevlavage A.R., Quackenbush J.,
 RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
 RA Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A.,
 RA Artiach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
 RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
 RA "Genomic sequence of a Lyme disease spirochete, Borrelia
 RT burgdorferi";
 RL Nature 390:580-586(1997).
 DR EMBL; AE001575; AAF07607.1; -; Genomic DNA.
 DR TIGR; BBO1; -;
 DR InterPro; IPR008505; DUF787.
 DR Pfam; PF05619; DUF787; 1.
 KW Complete proteome; Hypothetical protein; Plasmid.
 SQ SEQUENCE 370 AA; 42103 MW; 62FB9015CFEBBCD6 CRC64;
 Query Match 100.0%; Score 329; DB 2; Length 370;
 Best Local Similarity 100.0%; Pred. No. 6.6e-30;
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LAATPLVDDTVIGKLTAKINFYSLNETGLDGPAPKGVLAGGAIDEQFTYHYKNE 60
 Db 211 LAATPLVDDTVIGKLTAKINFYSLNETGLDGPAPKGVLAGGAIDEQFTYHYKNE 270
 QY 61 AIIEL 65
 Db 271 AIIEL 275
 RESULT 4
 Q9S0C8 BORBU PRELIMINARY; PRT; 370 AA.
 ID Q9S0C8 BORBU PRELIMINARY; PRT; 370 AA.
 AC Q9S0C8;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein BBM11.
 GN OrderedLocustNames=BBM11;
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OG Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
 OC Borrelia burgdorferi group.
 ON NCBI_TaxID=139;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 35210 / B31;
 RX MEDLINE=20138354; PubMed=10672174;
 RA Casjens S., Palmer N., van Vugt R., Huang W.M., Stevenson B., Rosa P.,
 RA Lathigra R., Sutton G.G., Peterson J.D., Dodson R.J., Haft D.H.,
 RA Hickey E.K., Gwinn M.L., White O., Fraser C.M.;
 RA "A bacterial genome in flux: the twelve linear and nine circular
 RT extrachromosomal DNAs in an infectious isolate of the Lyme disease
 RT spirochete Borrelia burgdorferi";
 RL Mol. Microbiol. 35:490-516(2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 35210 / B31;
 RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,

RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
 RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
 RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
 RA Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A.,
 RA Artach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
 RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.,
 RA "Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi."
 RL Nature 390:580-586(1997).
 DR EMBL; AE001578; AAF07566.1; -; Genomic_DNA.
 TIGR; BBN11; -.
 DR InterPro; IPR008505; DUF787.
 DR Pfam; PF05619; DUF787; 1.
 KW Complete proteome; Hypothetical protein; Plasmid.
 SQ SEQUENCE 370 AA; 42158 MW; 62FEE72EB97CF9C3 CRC64;

 Query Match 100.0%; Score 329; DB 2; Length 370;
 Best Local Similarity 100.0%; Pred. No. 6.6e-30;
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 LAATPLVDDTVIGKLTAKINFYSLNLTGDLGVPAFKEGVDLAGGAIDEQTYHYIKNE 60
 |||||
 DB 211 LAATPLVDDTVIGKLTAKINFYSLNLTGDLGVPAFKEGVDLAGGAIDEQTYHYIKNE 270
 |||||
 QY 61 AIIEL 65
 |||||
 DB 271 AIIEL 275
 |||||

 RESULT 5
 Q9S0F9 BORBU PRELIMINARY; PRT; 370 AA.
 AC Q9S0F9 BORBU PRELIMINARY; PRT; 370 AA.
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Hypothetical protein BBR11.
 GN OrderedLocusNames=BBR11;
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OG Plasmid cp32-4.
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
 OC Borrelia burgdorferi group.
 OX NCBI_TaxID=139;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 35210 / B31;
 RX MEDLINE=98065943; PubMed=10672174;
 RA Casjens S., Palmer N., van Vugt R., Huang W.M., Stevenson B., Rosa P.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
 RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
 RA Hickey E.K., Gwinn M.L., White O., Fraser C.M.;
 RA "A bacterial genome in flux: the twelve linear and nine circular
 RT extrachromosomal DNAs in an infectious isolate of the Lyme disease
 RT spirochete Borrelia burgdorferi."
 RL Mol. Microbiol. 35:490-516(2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 35210 / B31;
 RX MEDLINE=98065943; PubMed=10672174;
 RA Casjens S., Palmer N., van Vugt R., Huang W.M., Stevenson B., Rosa P.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
 RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
 RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
 RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
 RA Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A.,
 RA Artach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
 RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
 RA "Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi."
 RL Nature 390:580-586(1997).
 DR EMBL; AE001577; AAF07520.1; -; Genomic_DNA.
 TIGR; BBR11; -.
 DR InterPro; IPR008505; DUF787.
 DR Pfam; PF05619; DUF787; 1.

KW Complete proteome; Hypothetical protein; Plasmid.
 SQ SEQUENCE 370 AA; 42130 MW; BE37D9DB14A1B0BD CRC64;

 Query Match 100.0%; Score 329; DB 2; Length 370;
 Best Local Similarity 100.0%; Pred. No. 6.6e-30;
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 LAATPLVDDTVIGKLTAKINFYSLNLTGDLGVPAFKEGVDLAGGAIDEQTYHYIKNE 60
 |||||
 DB 211 LAATPLVDDTVIGKLTAKINFYSLNLTGDLGVPAFKEGVDLAGGAIDEQTYHYIKNE 270
 |||||
 QY 61 AIIEL 65
 |||||
 DB 271 AIIEL 275
 |||||

 RESULT 6
 Q9S051 BORBU PRELIMINARY; PRT; 362 AA.
 ID Q9S051 BORBU PRELIMINARY; PRT; 362 AA.
 AC Q9S051 BORBU PRELIMINARY; PRT; 362 AA.
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Hypothetical protein BBN11.
 GN OrderedLocusNames=BBN11;
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OG Plasmid cp32-9.
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
 OC Borrelia burgdorferi group.
 OX NCBI_TaxID=139;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 35210 / B31;
 RX MEDLINE=20138354; PubMed=10672174;
 RA Casjens S., Palmer N., van Vugt R., Huang W.M., Stevenson B., Rosa P.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
 RA Hickey E.K., Gwinn M.L., White O., Fraser C.M.;
 RA "A bacterial genome in flux: the twelve linear and nine circular
 RT extrachromosomal DNAs in an infectious isolate of the Lyme disease
 RT spirochete Borrelia burgdorferi."
 RL Mol. Microbiol. 35:490-516(2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 35210 / B31;
 RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
 RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
 RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
 RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
 RA Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A.,
 RA Artach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
 RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
 RA "Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi."
 RL Nature 390:580-586(1997).
 DR EMBL; AE001581; AAF07685.1; -; Genomic_DNA.
 TIGR; BBN11; -.
 DR InterPro; IPR008505; DUF787.
 DR Pfam; PF05619; DUF787; 1.
 KW Complete proteome; Hypothetical protein; Plasmid.
 SQ SEQUENCE 362 AA; 41367 MW; CD76E4F607088587 CRC64;

 Query Match 96.0%; Score 316; DB 2; Length 362;
 Best Local Similarity 95.4%; Pred. No. 2.1e-28;
 Matches 62; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

 QY 1 LAATPLVDDTVIGKLTAKINFYSLNLTGDLGVPAFKEGVDLAGGAIDEQTYHYIKNE 60
 |||||
 DB 211 LAATPLVDDTVIGKLTAKINFYSLNLTGDLGVPAFKEGVDLAGGAIDEQTYHYIKNE 270
 |||||
 QY 61 AIIEL 65
 |||||

```

Db      271 AIIEL 275
RESULT 7
Q9S003 BORBU
ID Q9S003 BORBU PRELIMINARY; PRT; 370 AA.
AC Q9S003;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein BBQ18.
GN OrderedLocNames=BBQ18;
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp56.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=139;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RC MEDLINE=20138354; PubMed=10672174;
RA Casjens S., Palmer N., van Vugt R., Huang W.M., Stevenson B., Rosa P.,
RA Lathigra R., Sutton G.G., Peterson J.D., Dodson R.J., Haft D.H.,
RA Hickey E.K., Gwinn M.L., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A.,
RA Artiach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi."
RL Nature 390:580-586 (1997).
RL EMBL; AB001584; AAF07741.1; -; Genomic_DNA.
DR TIGR; BBQ18; -.
DR InterPro: IPR008505; DUF787.
DR Pfam; PF05619; DUF787; 1.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 370 AA; 42150 MW; DAE039E8F8EBD8 CRC64;

Query Match 92.4%; Score 304; DB 2; Length 370;
Best Local Similarity 92.3%; Pred. No. 5.6e-27;
Matches 60; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LAATPLVDVTGKLTAKINFYSLNLTGDPAPKFGVDLAGAIDEQFTYHYKNE 60
Db      211 LAATPLVDVTGKLTAKINFYSLNLTGDPAPKFGVDLAGAIDEQFTYHYKNE 270
QY 61 AIIEL 65
Db      271 AIIEL 275
RESULT 8
Q9X268 BORBU
ID Q9X268 BORBU PRELIMINARY; PRT; 373 AA.
AC Q9X268;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=BGFP064;
OS Borrelia garinii PBI.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=290434;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBI;
RA Glockner G., Lehmann R., Romualdi A., Pradella S.,
RA Schloer-Spechtel U., Schilhabel M., Wilske B., Suhnel J., Platzner M.;
RT "Comparative analysis of the Borrelia garinii genome."
RT Nucleic Acids Res. 32:6038-6046 (2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBI;
RA Gloeckner G., Schilhabel M., Lehmann R., Platzner M.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY722919; AAU85914.1; -; Genomic_DNA.
DR InterPro: IPR008505; DUF787.
DR Pfam; PF05619; DUF787; 1.
KW Hypothetical protein.
SQ SEQUENCE 373 AA; 42512 MW; FC40FB96B757ED69 CRC64;

Query Match 90.6%; Score 298; DB 2; Length 373;
Best Local Similarity 89.2%; Pred. No. 2.8e-26;
Matches 58; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAATPLVDVTGKLTAKINFYSLNLTGDPAPKFGVDLAGAIDEQFTYHYKNE 60
Db      214 LAATPLVDVTGKLTAKINFYSLNLTGDPAPKFGVDLAGAIDEQFTYHYKNE 273
QY 61 AIIEL 65
Db      274 AIIEL 278
RESULT 9
Q9S039 BORBU
ID Q9S039 BORBU PRELIMINARY; PRT; 356 AA.
AC Q9S039;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein BBA46.
GN OrderedLocNames=BBA46;
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp54.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=139;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RC MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A.,
RA Artiach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi."
RL Nature 390:580-586 (1997).
RL EMBL; AB000790; AAC66278.1; -; Genomic_DNA.
DR PIR; F70212; F70212.
DR TIGR; BBA46; -.
DR InterPro: IPR008505; DUF787.
DR Pfam; PF05619; DUF787; 1.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 356 AA; 40359 MW; A11343C1F56ED474 CRC64;

Query Match 62.9%; Score 207; DB 2; Length 356;
Best Local Similarity 66.2%; Pred. No. 1.3e-15;

```

RA Lounseged H., Mungall K., Oliver K., Price C., Quail M.A., Sanders M.,
RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
RA Ma J.Y., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shaalsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzner M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuapa A.;
RT "The genome of the social amoeba Dictyostelium discoideum.";
CC Nature 0:0-0(2005).
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AF101000127; EAL64517.1; --; Genomic_DNA.
DR Hypochemical protein.
SQ SEQUENCE 250 AA; 29183 MW; 16D29CAF2BE5080F CRC64;

Query Match 24.6%; Score 81; DB 2; Length 250;
Best Local Similarity 48.6%; Pred. No. 0.5;
Matches 18; Conservative 5; Mismatches 12; Indels 2; Gaps 1;

QY 31 LDGVPAFKGVDLAGGAI--DSQFTYHYIKNEAIIEL 65
||| ||| ||| ||| ||| ||| : :
Db 150 LDGLQVKKKGVCLYGGSIDCDEDFTFYYICNLTSLGI 186

RESULT 12
Q97BQ1 THEVO PRELIMINARY; PRT; 402 AA.
ID Q97BQ1 THEVO PRELIMINARY;
AC Q97BQ1;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Aspartate aminotransferase.
GN OrderedLocustNames=TV0404; ORFNames=TVG0393535;
OS Thermoplasma volcanium.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=50339;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=GSS1 / DSM 4299 / JCM 9571;
RX MEDLINE=20570466; PubMed=1121031; DOI=10.1073/pnas.97.26.14257;
RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,
RA Kawashima-Ohya Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,
RA Nunohiba T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;
RT "Archaeal adaptation to higher temperatures revealed by genomic
sequence of Thermoplasma volcanium".
RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).
DR EMBL; BA000011; BAB95946.1; --; Genomic DNA.
DR GO; GO:0008483; F:transaminase activity; IEA.
DR CO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR004839; Aminotrans_I/II.
DR InterPro; IPR000529; Ribosomal_S6.
DR Pfam; PF00155; Aminotran_1_2; I.
DR PROSITE; PS01048; RIBOSOMAL_S6; UNKNOWN 1.
KW Aminotransferase; Complete proteome; Transferase.
SQ SEQUENCE 402 AA; 45684 MW; 1CE50D5AB95AB5DA CRC64;

Query Match 22.3%; Score 73.5; DB 2; Length 402;
Best Local Similarity 24.8%; Pred. No. 6.3;
Matches 21; Conservative 12; Mismatches 25; Indels 29; Gaps 2;

QY 5 PLVDDTIVIGKLR-----TAKINFYSLNLTG|-----DGVP 35
||:|:|||:|||| : :|:
Db 204 PIVEDNPYGQLRDGEFPVQSISMDKNNGNVILGTFSKWAPGLRGLGYVIAPREVIDKIN 263
||:|:|||:|||| : :|:
QY 36 AFKEGVDLAGGAIDEQFTYHYIKNEAI 62
||:|:|||:|||| : :|:
Db 264 LLKQGLDLASDSLSEIAYEYLKRGAI 290

RESULT 13
MECA LACLA

Lee K., Berry K.J., Lee C., Mueller J., Khouri H.M., Gill J.,
Utterback T.R., McDonald L.A., Feldblum T.V., Smith H.O.,
Venter J.C., Nealson K.H., Fraser C.M.;
"Genome sequence of the dissimilatory metal ion-reducing bacterium
Shewanella oneidensis";
Nat. Biotechnol. 20:1118-1123 (2002).
EMBL; AS015784; AAN56477.1; -; Genomic_DNA.
HSSP; P31224; IIWG.
TIGR; SO3484; -.
InterPro; IPR001036; Acrflvin_res.
Pfam; PF00873; ACR_tran; 1.
PRINTS; PR00702; ACRIFLAVINRP.
KW Complete proteome.
SQ SEQUENCE 1046 AA; 113938 MW; 49B17826F50EE5D CRC64;

Query Match 21.1%; Score 69.5; DB 2; Length 1046;
Best Local Similarity 31.2%; Pred. No.54;
Matches 20; Conservative 15; Mismatches 20; Indels 9; Gaps 3;

QY 8 DDTVGKLRTRAKI-----NFYSLLNETGLDGVPAPAF-KEGVDLAGGAIDEQTYHYIKNEA 61
:-|::|||:::-|-|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db 221 IDTTVMSTVKRLYYTPKDFYLIVRTASDGTPIYLKDVADVAGAQNENSTE---KSDG 277

QY 62 NYETL 65
:-|::|:
Db 278 IVNLT 281

RESULT 15
Q/QW/L_GIALA PRELIMINARY; PRT; 6858 AA.

ID Q7QWL1_GIALA PRELIMINARY; AC Q7QUW1;
AC AC
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DT GIU_561.28478.7902.
OS Giardia lamblia ATCC 50803.
OU Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
OX NCBI_TaxID=184922;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=WB C6;
RA Morrison H.G., McArthur A.G., Adam R.D., Alely S.B., Gillin F.D.,
RA Olsen G.J., Sogin M.L.;
RT "Draft sequence of the Giardia lamblia genome.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACB01000089; EAA38820.1; -; Genomic_DNA.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000264; Serum_albumin.
DR InterPro; IPR005829; Sug_transporter.
DR PROSITE; PS00212; ALBUMIN; UNKNOWN 1.
DR PROSITE; PS00615; C TYPE LECTIN 1; UNKNOWN 1.
DR PROSITE; PS00217; SUGAR TRANSPORT 2; UNKNOWN 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN 1.
SQ SEQUENCE 6858 AA; 767885 MW; 0B6D765D899303D7E CRC64;

Query Match 20.8%; Score 68.5; DB 2; Length 6858;
Best Local Similarity 28.8%; Pred. No.5.6e+02;
Matches 19; Conservative 8; Mismatches 22; Indels 17; Gaps 1;

QY 8 DDTVGKLRTRAKINFYSLLNETGLDGVPAPAKEGVDL-----AGGAIDE 50
:-|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db 1754 DISSIGAIRTPELFYLLNQTNLEGDPVFEPASTGSTSAVCRRHYEFFINALQSMTIDD 1813

QY 51 QFTVHY 56
:-|::|:
Db 1814 ESUYFY 1819

Search completed: January 24, 2006, 19:56:58
Job time : 15.9671 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:30:59 ; Search time 4.00711 Seconds
(without alignments)
1341.095 Million cell updates/sec

Title: US-10-688-058-18
Perfect score: 329
Sequence: 1 LAAPPLVDVTGKLTAKI.....GAIDEQFTYHYKNEAIIEI 65

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5 COMB pep.*
2: /cgn2_6/ptodata/1/iaa/6 COMB pep.*
3: /cgn2_6/ptodata/1/iaa/H COMB pep.*
4: /cgn2_6/ptodata/1/iaa/PTUS COMB pep.*
5: /cgn2_6/ptodata/1/iaa/RE COMB pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1 pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65.5	19.9	214	2	US-09-248-796A-17440
2	63	19.1	381	2	US-09-134-000C-5229
3	63	19.1	491	2	US-09-029-267-2
4	62	18.8	351	2	US-09-902-540-12944
5	61.5	18.7	302	2	US-09-710-279-2980
6	61.5	18.7	307	2	US-09-270-767-44911
7	61.5	18.7	327	2	US-09-489-039A-10690
8	61.5	18.7	350	2	US-09-134-001C-3383
9	61	18.5	371	2	US-09-134-000C-5173
10	60.5	18.4	612	2	US-09-538-092-775
11	60.5	18.4	645	2	US-09-252-991A-23253
12	59.5	18.1	137	2	US-09-710-279-530
13	58	17.6	647	2	US-09-252-991A-17460
14	58	17.6	1326	2	US-09-489-039A-7584
15	57	17.3	568	2	US-09-628-966-10
16	57	17.3	969	2	US-09-949-016-8059
17	56.5	17.2	274	2	US-09-248-796A-17624
18	56.5	17.2	5087	2	US-09-107-532A-5087
19	56.5	17.2	529	2	US-09-248-796A-15199
20	56.5	17.2	769	2	US-09-543-681A-7175
21	56	17.0	190	2	US-09-902-540-9718
22	56	17.0	202	2	US-09-252-991A-17985
23	56	17.0	331	2	US-09-248-796A-26542
24	56	17.0	1038	2	US-09-538-092-487
25	55.5	16.9	244	2	US-09-252-991A-29692
26	55.5	16.9	496	2	US-09-252-991A-26781
27	55	16.7	392	2	US-09-583-110-3805

28	55	16.7	410	2	US-09-107-433-5135	Sequence 5135, Ap
29	55	16.7	450	2	US-09-252-991A-25760	Sequence 25760, A
30	54.5	16.6	400	2	US-09-352-990-18	Sequence 18, Appl
31	54.5	16.6	413	2	US-10-360-101-249	Sequence 249, App
32	54.5	16.6	444	1	US-08-559-260-2	Sequence 2, Appli
33	54	16.4	216	2	US-09-270-767-45064	Sequence 45064, A
34	54	16.4	362	2	US-09-252-991A-24127	Sequence 24127, A
35	54	16.4	666	2	US-09-252-991A-32542	Sequence 32542, A
36	54	16.4	2736	2	US-09-252-991A-30227	Sequence 30227, A
37	53.5	16.3	187	2	US-09-673-395A-369	Sequence 369, App
38	53.5	16.3	519	2	US-09-134-000C-4563	Sequence 4563, Ap
39	53.5	16.3	627	2	US-09-328-352-4917	Sequence 4917, Ap
40	53.5	16.3	630	2	US-09-328-352-7722	Sequence 7722, Ap
41	53.5	16.3	1318	2	US-09-540-236-3623	Sequence 3623, Ap
42	53	16.1	117	2	US-09-540-236-2827	Sequence 2827, Ap
43	53	16.1	242	2	US-09-605-703B-148	Sequence 148, App
44	53	16.1	242	2	US-09-605-703B-150	Sequence 150, App
45	53	16.1	255	2	US-09-543-681A-6904	Sequence 6904, Ap

ALIGNMENTS

RESULT 1

US-09-248-796A-17440
Sequence 17440, Application US/09248796A
Patent No. 6747137

GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.132

CURRENT APPLICATION NUMBER: US/09/248,796A

CURRENT FILING DATE: 1999-02-12

PRIOR APPLICATION NUMBER: US 60/074,725

PRIOR FILING DATE: 1998-02-13

PRIOR APPLICATION NUMBER: US 60/096,409

PRIOR FILING DATE: 1998-08-13

NUMBER OF SEQ ID NOS: 28208

SEQ ID NO 17440

LENGTH: 214

TYPE: PRT

ORGANISM: Candida albicans

US-09-248-796A-17440

Query Match 19.9%; Score 65.5; DB 2; Length 214;

Best Local Similarity 27.5%; Pred. No. 0.65;

Matches 19; Conservative 11; Mismatches 22; Indels 17; Gaps 3;

QY 12 IOKLTAKINFYSLNETGLGVPAF-----KEGVLAGGAIDQFTYHY---56

Db 44 VAKQAARKVDNFKNVGLGTGTASSLTAFKKRDEAKKEFIDLSQSOTDIDFN-HYRS 102

QY 57 -IKNEAIE 64

Db 103 ILKNSKVD 111

RESULT 2

US-09-134-000C-5229

Sequence 5229, Application US/09134000C

Patent No. 6617156

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 032796-032

CURRENT APPLICATION NUMBER: US/09/134,000C

CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/055,778

PRIOR FILING DATE: 1997-08-15

NUMBER OF SEQ ID NOS: 6812

SOFTWARE: Patentin version 3.1

```

; SEQ ID NO 5229
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5229

Query Match      19.1%; Score 63; DB 2; Length 381;
Best Local Similarity 34.4%; Pred. No. 3.1;
Matches 22; Conservative 10; Mismatches 22; Indels 10; Gaps 3;

QY 10 TWIGKLTAKINFYSLNETGLD-----GVPAPKEGVDLGAGAIDE-QFTYHYIKNEA 61
Db 185 TVVAGRAINEMN--SNINVTYDKKIIDVNDVVELTEGIDIIVCAIDPPFLIHRIVNEA 242

QY 62 ITEL 65
Db 243 IVKV 246

RESULT 3
US-09-029-267-2
; Sequence 2, Application US/09029267
; Patent No. 6107057
; GENERAL INFORMATION:
; APPLICANT: Crawford, Kenneth
; APPLICANT: Zaror, Isabel
; APPLICANT: Innis, Michael
; TITLE OF INVENTION: Pichia Secretory Leader for Protein
; TITLE OF INVENTION: Expression
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: United States
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/029,267
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Chung, Ling-Fong
; REGISTRATION NUMBER: 36,482
; REFERENCE/DOCKET NUMBER: 1165.100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2704
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 491 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-029-267-2

Query Match      19.1%; Score 63; DB 2; Length 491;
Best Local Similarity 25.4%; Pred. No. 4.3;
Matches 16; Conservative 11; Mismatches 16; Indels 20; Gaps 2;

QY 9 DTVIG-----KLRTAKINFYSLNETGLDGVPAKKEGVDLGAGAIDEQFTYHYIKNEAI 62
Db 278 DNIVGVMVLQRIINTENLNFSLTNEVTKNGVP-----DHFYVEYAKNDPM 323

QY 63 IEL 65
Db 324 IQI 326

```

```

RESULT 4
US-09-902-540-12944
; Sequence 12944, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wisegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12944
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12944

Query Match      18.8%; Score 62; DB 2; Length 351;
Best Local Similarity 37.0%; Pred. No. 3.8;
Matches 20; Conservative 6; Mismatches 14; Indels 14; Gaps 4;

QY 7 VDDTVIGKLTAKINFYSLNETGLDGVPAKKEGVDLGAGAIDEQFT---YHYI 57
Db 290 VDLTYLGHIIH---SYTSFANA---GIPAFISG---GGGAIPETFDGVRHYL 332

RESULT 5
US-09-710-279-2980
; Sequence 2980, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2980
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-2980

Query Match      18.7%; Score 61.5; DB 2; Length 302;
Best Local Similarity 32.1%; Pred. No. 3.7;
Matches 18; Conservative 10; Mismatches 17; Indels 11; Gaps 2;

QY 4 TPLVDDTVIGKLTAKINFYSLNETGLDGVPAKKEGVDLGAGAIDEQFTYHYIKN 59
Db 160 TPLAEFNIYCDPEAAQIVFNS-----GLPLTWIGLDLAREAL---FTTHFVKD 204

RESULT 6
US-09-270-767-44911
; Sequence 44911, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094

```



```

; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44911
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-44911

Query Match      18.7%   Score 61.5; DB 2; Length 307;
Best Local Similarity 30.4%; Pred.No.3.8;
Matches 14; Conservative 9; Mismatches 20; Indels 3; Gaps 1;

QY      20 INFYSLINETGLDGPAPKEGVDLGAGIDQFTVHYIKNEAIIEL 65
      : : : : : : : : : : : : : : : : : : : : : :
Db      193 LDFEFLNNINVTALTFF---YHAGASIDQOTLQHVAKTVAMVNL 235

RESULT 7
US-09-489-039A-10690
; Sequence 10690, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10690
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10690

Query Match      18.7%   Score 61.5; DB 2; Length 327;
Best Local Similarity 37.2%; Pred.No.4.1;
Matches 16; Conservative 8; Mismatches 16; Indels 3; Gaps 1;

QY      24 SLINETGLDGPAPKEGVDLA---GGAIDQFTVHYIKNEAII 63
      : : : : : : : : : : : : : : : : : : : : : :
Db      153 SLTVLTGNDVNVLQAGIDLAIFYDDAPSSQLSHFLMDEAIV 195

RESULT 8
US-09-134-001C-3383
; Sequence 3383, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3383
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3383

Query Match      18.7%   Score 61.5; DB 2; Length 350;

```

RESULT 11
 US-09-252-991A-25253
 ; Sequence 25253, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 25253
 ; LENGTH: 645
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-25253
 Query Match 18.4%; Score 60.5; DB 2; Length 645;
 Best Local Similarity 36.2%; Pred. No. 14;
 Matches 17; Conservative 8; Mismatches 17; Indels 5; Gaps 2;
 QY 11 VIGKLRTA-KINFYSLNETGLDGVPAKFGVDLAGGAIDQFTYHY 56
 DB 114 VTNKLEATSIHWHGIILPYQMDGVP----GISFNGIAPGETTYRF 156
 RESULT 12
 US-09-710-279-530
 ; Sequence 530, Application US/09710279
 ; Patent No. 6703492
 ; GENERAL INFORMATION:
 ; APPLICANT: KIMMERLY, WILLIAM JOHN
 ; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
 ; FILE REFERENCE: PU3480US
 ; CURRENT APPLICATION NUMBER: US/09/710,279
 ; CURRENT FILING DATE: 2000-11-09
 ; PRIOR APPLICATION NUMBER: 60/164,258
 ; PRIOR FILING DATE: 1999-11-09
 ; NUMBER OF SEQ ID NOS: 4472
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 530
 ; LENGTH: 137
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic
 ; OTHER INFORMATION: amino acid sequence
 US-09-710-279-530
 Query Match 18.1%; Score 59.5; DB 2; Length 137;
 Best Local Similarity 33.9%; Pred. No. 2.5;
 Matches 21; Conservative 9; Mismatches 23; Indels 9; Gaps 3;
 QY 7 VDDTVIGKLRKAKINFYSLNETGLDGVPAKFGVDLAGGAI-DEQFTYHYIKNEAI 62
 DB 16 VNDGSIYGITKDLVRDYLVMQSTGLKD----KNGVEIYEGDIIPEDESFQYPPF-DDEAI 70
 QY 63 IE 64
 DB 71 VE 72
 RESULT 13
 US-09-252-991A-17460
 ; Sequence 17460, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 17460
 ; LENGTH: 647
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-17460
 Query Match 17.6%; Score 58; DB 2; Length 647;
 Best Local Similarity 40.0%; Pred. No. 31;
 Matches 12; Conservative 6; Mismatches 12; Indels 0; Gaps 0;
 QY 25 LLNETGLDGVPAKFGVDLAGGAIDQFTY 54
 DB 337 LLHEAGVDNLLTFSIGFEDAGGERGDFQY 366
 RESULT 14
 US-09-489-039A-7584
 ; Sequence 7584, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 7584
 ; LENGTH: 1326
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-7584
 Query Match 17.6%; Score 58; DB 2; Length 1326;
 Best Local Similarity 33.9%; Pred. No. 79;
 Matches 21; Conservative 6; Mismatches 19; Indels 16; Gaps 3;
 QY 13 GKLRKAKINFYSLNET-GLDGVPAKFGVDLAGGAIDQFTYHYIKNEAI 62
 DB 218 GLLFTGKL--VSTHNETSLRSLNRIGSGEPLIRKGVDMARLUMGQF-----VTGETI 271
 QY 63 IE 64
 DB 272 AE 273
 RESULT 15
 US-09-628-966-10
 ; Sequence 10, Application US/09628966
 ; Patent No. 6596527
 ; GENERAL INFORMATION:
 ; APPLICANT: YEH, EDWARD T.H.
 ; APPLICANT: GONG, LIMIN
 ; TITLE OF INVENTION: COMPOSITION AND METHODS RELATING TO SENP1 - A
 ; FILE REFERENCE: UTSH-245U1
 ; CURRENT APPLICATION NUMBER: US/09/628,966
 ; CURRENT FILING DATE: 2000-07-31
 ; PRIOR APPLICATION NUMBER: 60/146,774
 ; PRIOR FILING DATE: 1999-07-31

This Page Blank (uspto)

Result No.	Score	Query Match	Length	DB	ID	Description	
1	329	100.0	65	5	US-10-688-058-18	Sequence 18, Appl	
2	329	100.0	160	5	US-10-688-058-62	Sequence 62, Appl	
3	329	100.0	370	5	US-10-688-058-20	Sequence 20, Appl	
4	329	100.0	370	5	US-10-688-058-64	Sequence 64, Appl	
5	67	20.4	182	4	US-10-424-599-151311	Sequence 151311, Appl	
6	67	20.4	188	5	US-10-739-930-10002	Sequence 10002, A	
7	65	19.8	445	5	US-10-732-923-2799	Sequence 2799, Ap	
8	64.5	19.6	601	4	US-10-282-122A-61809	Sequence 61809, A	
9	64	19.5	100	4	US-10-767-701-32620	Sequence 32620, A	
10	64	19.5	124	4	US-10-425-115-210704	Sequence 210704, A	
11	64	19.5	138	4	US-10-425-114-61645	Sequence 61645, A	
12	64	19.5	148	5	US-10-739-930-8484	Sequence 8484, Ap	
13	64	19.5	182	4	US-10-437-963-187208	Sequence 187208, A	
14	64	19.5	193	4	US-10-425-114-48053	Sequence 48053, A	
15	64	19.5	193	4	US-10-425-114-57200	Sequence 57200, A	
16	64	19.5	193	4	US-10-425-114-60717	Sequence 60717, A	
17	64	19.5	183	4	US-10-425-114-70412	Sequence 70412, A	
18	63.5	19.3	301	4	US-10-369-493-576	Sequence 576, App	
19	63.5	19.3	444	4	US-10-424-599-271469	Sequence 271469, A	
20	63	19.1	380	4	US-10-282-122A-57349	Sequence 57349, A	
21	63	19.1	489	4	US-10-424-599-237607	Sequence 237607, A	
22	63	19.1	491	4	US-10-179-046-2	Sequence 2, Appli	
23	62.5	19.0	402	4	US-10-369-493-18252	Sequence 18252, A	
24	62.5	19.0	422	4	US-10-437-963-103770	Sequence 103770, A	
25	62.5	19.0	1017	4	US-10-369-493-20166	Sequence 20166, A	
26	62	18.8	162	4	US-10-767-701-54228	Sequence 54228, A	
27	62	18.8	750	4	US-10-171-404A-2	Sequence 2, Appli	

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-688-058-62

Query Match          100.0%; Score 329; DB 5; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.5e-36;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LAATPLVDDTVIGKLTAKINFYSLNETGLDGVPAFKEGVDLAGGAIDQFTYHYIKNE 60
   |||||||
Db 1 LAATPLVDDTVIGKLTAKINFYSLNETGLDGVPAFKEGVDLAGGAIDQFTYHYIKNE 60
   |||||||

QY 61 AIIEL 65
   |||||
Db 271 AIIEL 275
   |||||

RESULT 3
US-10-424-599-151311
; Sequence 151311, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 151311
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_107657C.1.pep
US-10-424-599-151311

Query Match          20.4%; Score 67; DB 4; Length 182;
Best Local Similarity 35.0%; Pred. No. 1.3;
Matches 21; Conservative 8; Mismatches 25; Indels 6; Gaps 2;

QY 6 LVDDTVIGKLTAKINFYSLNETGLDGVPAFKEGVDLAG----GAIDQFTYHYIKN 59
   :|| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 91 VVDSDDVRLVIAKEEFHAILEEBELKGAVVLIIFANKQDLFCALDDAAVTEALELHKIKN 150
   :|| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

RESULT 6
US-10-739-930-10002
; Sequence 10002, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 10002
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) : (188)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-C235_53.p
US-10-739-930-10002
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-688-058-62

Query Match          100.0%; Score 329; DB 5; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.5e-36;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LAATPLVDDTVIGKLTAKINFYSLNETGLDGVPAFKEGVDLAGGAIDQFTYHYIKNE 60
   |||||||
Db 1 LAATPLVDDTVIGKLTAKINFYSLNETGLDGVPAFKEGVDLAGGAIDQFTYHYIKNE 60
   |||||||

QY 61 AIIEL 65
   |||||
Db 271 AIIEL 275
   |||||

RESULT 3
US-10-688-058-20
; Sequence 20, Application US/10688058
; Publication No. US20050058661A1
; GENERAL INFORMATION:
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: HALE, KATHERINE S.
; APPLICANT: JOHNSTON, STEPHEN A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF THE GENUS
; TITLE OF INVENTION: BORRELIA
; FILE REFERENCE: MCRO:003US
; CURRENT APPLICATION NUMBER: US/10/688,058
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-688-058-20

Query Match          100.0%; Score 329; DB 5; Length 370;
Best Local Similarity 100.0%; Pred. No. 4.6e-36;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LAATPLVDDTVIGKLTAKINFYSLNETGLDGVPAFKEGVDLAGGAIDQFTYHYIKNE 60
   |||||||
Db 211 LAATPLVDDTVIGKLTAKINFYSLNETGLDGVPAFKEGVDLAGGAIDQFTYHYIKNE 270
   |||||||

QY 61 AIIEL 65
   |||||
Db 271 AIIEL 275
   |||||

RESULT 4
US-10-688-058-64
; Sequence 64, Application US/10688058
; Publication No. US20050058661A1
; GENERAL INFORMATION:
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: HALE, KATHERINE S.
; APPLICANT: JOHNSTON, STEPHEN A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF THE GENUS
; TITLE OF INVENTION: BORRELIA
; FILE REFERENCE: MCRO:003US
; CURRENT APPLICATION NUMBER: US/10/688,058
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 370
```

Query Match 20.4%; Score 67; DB 5; Length 188;
Best Local Similarity 36.7%; Pred. No. 1.3;
Matches 22; Conservative 7; Mismatches 25; Indels 6; Gaps 2;

Qy 6 LVDDTVIGKLTAKINFYSLNETGLDG--VPAFKEGVLDAG---GAIDQFTYHYIKN 59
Db 97 VVDSDDTLRVTAKEEFHSLLEDELKGVVLYVANKODLFGALDDAAITSESLHLIKS 156

RESULT 7
US-10-732-923-2799
; Sequence 2799, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 2799
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Asterina pectinifera
US-10-732-923-2799

Query Match 19.8%; Score 65; DB 5; Length 445;
Best Local Similarity 26.3%; Pred. No. 7.9;
Matches 15; Conservative 12; Mismatches 18; Indels 12; Gaps 1;

Qy 2 AATPLVDDTVIGKLTAKINFYSLNETGLDGVPAPKEGVLDAGGAIDQFTYHYIK 58
Db 156 SASPMVLDTLSSENTAQTIEDIDNSDGVFGVPEYAEI-----YELR 200

RESULT 8
US-10-282-122A-61809
; Sequence 61809, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Lianguo
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patentin version 3.1
SEQ ID NO 61809
LENGTH: 601
TYPE: PRT
ORGANISM: Mycobacterium avium
US-10-282-122A-61809

Query Match 19.6%; Score 64.5; DB 4; Length 601;
Best Local Similarity 32.7%; Pred. No. 14;
Matches 17; Conservative 8; Mismatches 24; Indels 3; Gaps 1;

Qy 6 LVDDTVIGKLTAKIN---FYSLINETGLDGVPAFKEGVLDAGGAIDQFTY 54
Db 257 LVADVPVGCLLSGVSSLIWGLAEAGOHGLSTFSIGFESAGVGKDEPQY 308

RESULT 9
US-10-767-701-32620
; Sequence 32620, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 32620
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C14541_1.pep
US-10-767-701-32620

Query Match 19.5%; Score 64; DB 4; Length 100;
Best Local Similarity 35.0%; Pred. No. 1.5;
Matches 21; Conservative 8; Mismatches 25; Indels 6; Gaps 2;

Qy 6 LVDDTVIGKLTAKINFYSLNETGLDG--VPAFKEGVLDAG---GAIDQFTYHYIKN 59
Db 9 VVDSDDTLRVTAKEEFHSLLEDELKGVVLYVANKODLFGALDDAAITSESLHLIKS 68

RESULT 10
US-10-425-115-210704
; Sequence 210704, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 210704
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Zea mays

Search completed: January 24, 2006, 20:54:43
Job time : 12.4241 secs

This Page Blank (uspto)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	61.5	18.7	302	6	US-10-793-626-2980	Sequence 2980, Ap
2	59.5	18.1	137	6	US-10-793-626-530	Sequence 530, App
3	59	17.9	227	6	US-10-510-386-60	Sequence 60, Appl
4	52	15.8	195	7	US-11-055-822-914	Sequence 914, App
5	52	15.8	314	6	US-10-485-517-230	Sequence 230, App
6	52	15.8	762	7	US-11-055-822-912	Sequence 912, App
7	51.5	15.7	959	6	US-10-467-962B-4	Sequence 4, Appli
8	51	15.5	333	6	US-10-467-657-6450	Sequence 6450, Ap
9	51	15.5	437	6	US-10-467-657-5526	Sequence 5526, Ap
10	50	15.2	168	7	US-11-108-172-198	Sequence 198, App
11	50	15.2	412	7	US-11-055-822-1086	Sequence 1086, Ap
12	50	15.2	520	6	US-10-878-556A-38	Sequence 38, Appl
13	49.5	15.0	91	7	US-11-073-347-89	Sequence 89, Appl
14	49.5	15.0	147	7	US-11-073-347-90	Sequence 90, Appl
15	49.5	15.0	308	6	US-10-467-657-4192	Sequence 4192, Ap
16	49.5	15.0	530	6	US-10-523-038-41	Sequence 41, Appl
17	49.5	15.0	738	6	US-10-523-038-39	Sequence 39, Appl
18	49	14.9	298	6	US-10-467-657-4362	Sequence 4362, Ap
19	49	14.9	447	7	US-11-024-959-286	Sequence 286, App
20	49	14.9	465	6	US-10-793-626-1676	Sequence 1676, Ap
21	49	14.9	622	7	US-11-012-443-185	Sequence 185, App
22	48.5	14.7	129	6	US-10-467-657-4354	Sequence 4354, Ap
23	48.5	14.7	205	7	US-11-134-795-27	Sequence 27, Appl
24	48.5	14.7	345	6	US-10-467-657-252	Sequence 252, App
25	48.5	14.7	345	6	US-10-467-657-3086	Sequence 3086, Ap

US-11-055-822-912

; Sequence 912, Application US/11055822
; Publication No. US20050260707A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
; CURRENT APPLICATION NUMBER: US/11/055,822
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: 09/606,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142,101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148,613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187,970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931415.2
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931418.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1158
; SEQ ID NO 912
; LENGTH: 762
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-055-822-912

Query Match 15.8%; Score 52; DB 7; Length 762;
Best Local Similarity 31.7%; Pred. No. 70;
Matches 19; Conservative 8; Mismatches 19; Indels 14; Gaps 3;
QY 2 AATPLVDTVIG--KLTAKINFYS-----LLNETGLDGVPAFKEGVDLAGGAIDEQ 51
DB 190 AGNPLVNALCVGTLKVEDKLAFASGTGNKVLFGSRTGLDGI-----GGVSVLGSASPEE 245

RESULT 7

US-10-467-962B-4
; Sequence 4, Application US/10467962B
; Publication No. US20050246784A1
; GENERAL INFORMATION:
; APPLICANT: Plesch, Gunnar
; APPLICANT: Blau, Astrid
; APPLICANT: Daeschner, Klaus
; APPLICANT: Klein, Mathieu
; TITLE OF INVENTION: Identification of Herbicidally Active Substances
; FILE REFERENCE: 2000 857
; CURRENT APPLICATION NUMBER: US/10/467,962B
; CURRENT FILING DATE: 2003-08-14
; PRIOR APPLICATION NUMBER: PCT/EP02/01466
; PRIOR FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 4
; LENGTH: 959
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-467-962B-4

Query Match 15.7%; Score 51.5; DB 6; Length 959;
Best Local Similarity 36.1%; Pred. No. 1.1e+02;
Matches 13; Conservative 6; Mismatches 16; Indels 1; Gaps 1;

QY 6 LVDDTVIGKLTAKINFYSLLNETGLDGVPAFKEGV 41
DB 209 LADDNGLKTLQSLITLYLTLLCQ-GFDGTPWVKKAI 243

RESULT 8

US-10-467-657-6450
; Sequence 6450, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6450
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6450

Query Match 15.5%; Score 51; DB 6; Length 333;
Best Local Similarity 27.7%; Pred. No. 32;
Matches 18; Conservative 12; Mismatches 27; Indels 8; Gaps 2;

QY 6 LVDDTVIGKLTAKINFYSLLNETGLDGV--PAFKEGVDLAGGAIDEQTYHYI-----57
DB 146 LADNIIAEKVRQQAQMNSRVSEAEIDAFLEBQAQKQGITLPEGAPLROYRAQHILIKADS 205

QY 58 KNEAI 62
DB 206 KNAAV 210

RESULT 9

US-10-467-657-5526
; Sequence 5526, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5526
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5526

Query Match 15.5%; Score 51; DB 6; Length 437;
Best Local Similarity 36.7%; Pred. No. 45;

Matches 11; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

Qy 23 YSLNETGLDGVPAFKEGVLAGGAIDEQF 52
Db 332 YEVLEKLEDSFKFLAMELEQIALKDPF 361

RESULT 10

US-11-108-172-198
; Sequence 198, Application US/11108172
; Publication No. US20050260177A1

; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun

; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather

; APPLICANT: Benson, Davin R.
; APPLICANT: Meagher, Madeleine Joy

; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong

; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.

; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aljun

; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.

; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.

; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS

; FILE REFERENCE: 210121.471C15
; CURRENT APPLICATION NUMBER: US/11/108,172

; PRIOR FILING DATE: 2005-04-15
; PRIOR FILING DATE: 2001-12-19

; PRIOR FILING DATE: 2001-08-03
; PRIOR FILING DATE: 2001-08-03

; PRIOR FILING DATE: 2001-04-10
; PRIOR FILING DATE: 2000-08-28

; PRIOR FILING DATE: 2000-06-29
; PRIOR FILING DATE: 2000-05-19

; PRIOR FILING DATE: 2000-03-06
; PRIOR FILING DATE: 2000-02-15

; PRIOR FILING DATE: 2000-01-10
; PRIOR FILING DATE: 2000-01-10

; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 1130

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 198

; TYPE: PRT
; ORGANISM: Homo sapiens

US-11-108-172-198

Query Match 15.2%; Score 50; DB 7; Length 168;
Best Local Similarity 32.4%; Pred. No. 17;

Matches 11; Conservative 7; Mismatches 12; Indels 4; Gaps 1;

Qy 9 DTVIGKLTAKINFYSLNETGLDGVPAFKEGVD 42
Db 133 ETIIDKSKVKTNLQLFESGNWTDI----EGID 162

RESULT 11

US-11-055-822-1086

Query Match 15.2%; Score 50; DB 6; Length 520;
Best Local Similarity 32.4%; Pred. No. 77;

Matches 11; Conservative 10; Mismatches 28; Indels 2; Gaps 2;

Qy 7 VDDTVIGKLTAKINFYSLNETGLDGVPAFKEGVDLA-GGAIDEQFYHYIKNEAI 62
Db 201 IDDTPLDPERDYV-AASLYLQSGNEGMTALTRGTAPAQGTGIVDVQSTIGYLSNNV 256

RESULT 12
US-10-878-556A-38

; Sequence 38, Application US/10878556A
; Publication No. US20050266399A1

; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.

; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21762

; CURRENT APPLICATION NUMBER: US/10/878,556A
; CURRENT FILING DATE: 2004-06-28

; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1

; Sequence 1086, Application US/11055822
; Publication No. US20050260707A1

; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus

; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig

; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor

; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; FILE REFERENCE: BGI-121PCN

; CURRENT APPLICATION NUMBER: US/11/055,822
; CURRENT FILING DATE: 2005-02-11

; PRIOR FILING DATE: 2000-06-23
; PRIOR FILING DATE: 1999-06-25

; PRIOR FILING DATE: 1999-07-02
; PRIOR FILING DATE: 1999-07-02

; PRIOR FILING DATE: 1999-08-12
; PRIOR FILING DATE: 2000-03-09

; PRIOR FILING DATE: 1999-07-01
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:21:03 ; Search time 82.6952 Seconds
(without alignments)
1966.133 Million cell updates/sec

Title: US-10-688-058-20
Perfect score: 1889
Sequence: 1 MPODTISVLLDSRIQASRP.....AVSLVITQDIVYQNSLSA 370

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq 21:.*
1: geneseqp1980s:.*
2: geneseqp1990s:.*
3: geneseqp2000s:.*
4: geneseqp2001s:.*
5: geneseqp2002s:.*
6: geneseqp2003as:.*
7: geneseqp2003bs:.*
8: geneseqp2004s:.*
9: geneseqp2005s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1889	100.0	370	ADU98761	Adu98761 Borrelia
2	1837	97.2	370	ADU98805	Adu98805 Borrelia
3	788	41.7	160	ADU98803	Adu98803 Borrelia
4	392	20.8	79	AAV83290	Aay83290 Borrelia
5	329	17.4	65	ADU98759	Adu98759 Borrelia
6	137.5	7.3	1633	ABU42513	Abu42513 Protein e
7	134.5	7.1	2295	AAU18180	Aab18180 Plasmid
8	132.5	7.0	1155	AAU82343	Ag82343 S. epider
9	130.5	6.9	1802	AAV83170	Aay83170 Cell wall
10	130.5	6.9	1802	AAV70119	Aay70119 Staph. ep
11	128.5	6.8	1979	AAU18171	Aab18171 Plasmid
12	120.5	6.4	575	ADU98188	Ad98188 Staphyloc
13	120.5	6.4	598	ABU26292	Abu26292 Protein e
14	119	6.3	1178	ABU44244	Abu44244 Protein e
15	119	6.3	1331	ABU35503	Abu35503 Protein e
16	119	6.3	1411	AAV31948	Aay31948 Plasmid
17	117.5	6.2	5005	ABU48947	Abu48947 Protein e
18	117	6.2	1156	ADU17390	Adn17390 Bacterial
19	117	6.2	1417	AAV31947	Aay31947 Plasmid
20	117	6.2	1471	ADP25445	Adp25445 Plasmid
21	116.5	6.2	1191	ABU24124	Abu24124 Protein e
22	116	6.1	524	ADU4502	Adf4502 Bacterial
23	115.5	6.1	439	ADI63080	Adi63080 Human apo
24	115.5	6.1	720	ABB77611	ABE77611 ANEPV RNA

25	115.5	6.1	1904	8	ADP25424	Adp25424 Plasmid
26	115	6.1	590	6	ABU24573	Abu24573 Protein e
27	115	6.1	4134	2	AAV31946	Aay31946 Plasmid
28	114.5	6.1	1100	3	AAU18213	Aab18213 Plasmid
29	113.5	6.0	596	6	ABU24511	Abu24511 Protein e
30	113.5	6.0	1516	3	AAU18195	Aab18195 Plasmid
31	113.5	6.0	1781	8	ADP25447	Adp25447 Plasmid
32	113.5	6.0	3029	8	ADP25433	Adp25433 Plasmid
33	113	6.0	1346	3	AAU18236	Aab18236 Plasmid
34	113	6.0	1805	6	ABU35589	Abu35589 Protein e
35	113	6.0	1805	7	ABO23572	ABO23572 Mycoplasma
36	112.5	6.0	326	5	ABU51677	Abu51677 Helicobac
37	112.5	6.0	382	2	AAU06932	Aaw06932 Candi locu
38	112.5	6.0	1285	5	ABP73331	Abp73331 Candida a
39	112	5.9	1254	2	AAU07503	Aaw07503 Merozoite
40	112	5.9	1254	2	AAW24575	Aaw24575 Merozoite
41	112	5.9	1881	5	ABP73809	Abp73809 Candida a
42	111.5	5.9	1172	5	ABP73666	Abp73666 Candida a
43	111	5.9	1087	2	AAV19935	Aay19935 B. burgdo
44	111	5.9	1099	2	AAW40538	Aaw40538 Mutant C-
45	111	5.9	1119	2	AAV19934	Aay19934 B. burgdo

ALIGNMENTS

RESULT 1
ADU98761
ID ADU98761 standard; protein; 370 AA.
XX ADU98761;
AC ADU98761;
XX
DT 24-FEB-2005 (first entry)
XX
DE Borrelia burgdorferi antigenic polypeptide seqid 20.
XX
KW antibacterial; vaccine; immune stimulation; immunity; antigen;
KW DNA library.
XX
OS Borrelia burgdorferi.
XX
FN WO2004103269-A2.
XX
PD 02-DEC-2004.
XX
PF 17-OCT-2003; 2003WO-US033056.
XX
PR 18-OCT-2002; 2002US-0419401P.
XX
PA (MACR-) MACROGENICS INC.
XX (TEXA) UNIV TEXAS SYSTEM.
XX Sykes KF, Hale KS, Johnston SA;
XX WPI; 2004-834155/82.
XX N-PSDB; ADU98760.
XX
PT Immunizing a subject against Borrelia burgdorferi infection comprises providing to the subject at least one Borrelia antigen or its fragment.
XX
Claim 27; SEQ ID NO 20; 121pp; English.
XX
CC The invention describes a method of immunizing a subject comprising providing to the subject at least one Borrelia antigen or its fragment to induce an immune response. Also described are: an isolated polynucleotide comprising a sequence having at least 17 contiguous nucleotides in common with a sequence not given in the specification; an isolated polypeptide having at least 5 consecutive amino acids of the sequence not given in the specification; a vaccine composition comprising at least one Borrelia antigen or at least one polynucleotide encoding a Borrelia antigen; screening for at least one test polypeptide or test polynucleotide encoding a polypeptide for an ability to produce an immune response; preparing a vaccine; vaccinating a subject; treating a subject infected

CC with a pathogen; passive immunization; vaccination; detecting Herpesvirus
 CC ; and detecting antibodies to a herpesvirus. The method is useful in
 CC immunizing a subject against Borrelia burgdorferi infection. This is the
 CC amino acids sequence of a Borrelia burgdorferi polypeptide that may be
 CC useful in immunizing a subject against infection by a member of the
 CC Borrelia species.
 XX
 SQ Sequence 370 AA;

Query Match 100.0%; Score 1889; DB 8; Length 370;
 Best Local Similarity 100.0%; Pred. No. 5.5e-154; Mismatches 0; Indels 0; Gaps 0;
 Matches 370; Conservative 0;

QY 1 MPQDTISVSLDSRIQASRPNNYPLVYKTAKIKVNDAAANYKILNLTNNYKQIETL 60
 DB 1 MPQDTISVSLDSRIQASRPNNYPLVYKTAKIKVNDAAANYKILNLTNNYKQIETL 60
 QY 61 EKDNNGQDQFGKTKLLKTAMSNFNSSESLKSADLFYKDKPEELKYLKVRHTFV 120
 DB 61 EKDNNGQDQFGKTKLLKTAMSNFNSSESLKSADLFYKDKPEELKYLKVRHTFV 120
 QY 121 VLINTEGNSDGLKIYKDDYDKFTPSIFFVSTKEQEIKELFKDKGNTKERNIAVYS 180
 DB 121 VLINTEGNSDGLKIYKDDYDKFTPSIFFVSTKEQEIKELFKDKGNTKERNIAVYS 180
 QY 181 NKNLNLHLKFIQYLHQASIFHAVNPYGMPLAATPLVDDTVIGKLTAKINFYSLNETG 240
 DB 181 NKNLNLHLKFIQYLHQASIFHAVNPYGMPLAATPLVDDTVIGKLTAKINFYSLNETG 240
 QY 241 LDGVPAPFKEGVDLAGGAIDQFTYHYIKNEAIIELIRIWNKNRQNSKLSALQLSGARDN 300
 DB 241 LDGVPAPFKEGVDLAGGAIDQFTYHYIKNEAIIELIRIWNKNRQNSKLSALQLSGARDN 300
 QY 301 AYTSAEICLLKRFVDRGLIIQYKNLSLTSPTPQLKLELSVNITNFSINAVSLVITTD 360
 DB 301 AYTSAEICLLKRFVDRGLIIQYKNLSLTSPTPQLKLELSVNITNFSINAVSLVITTD 360
 QY 361 IVDYQNSLSA 370
 DB 361 IVDYQNSLSA 370

RESULT 2
 ADU98805
 ID ADU98805 standard; protein; 370 AA.
 AC ADU98805;
 XX
 XX 24-FEB-2005 (first entry)
 DT
 DE Borrelia burgdorferi antigenic polypeptide seqid 64.
 XX
 XX antibacterial; vaccine; immune stimulation; immunity; antigen;
 KW DNA library.
 XX
 XX Borrelia burgdorferi.
 OS
 XX WO2004103269-A2.
 PN
 XX 02-DEC-2004.
 PD
 XX 17-OCT-2003; 2003WO-US033056.
 PF
 XX 18-OCT-2002; 2002US-0419401P.
 PR
 XX (MACR-) MACROGENICS INC.
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX
 XX Sykes KF, Hale KS, Johnston SA;
 PI WPI; 2004-834155/82.
 DR N-PSDB; ADU98804.
 XX

PT Immunizing a subject against Borrelia burgdorferi infection comprises
 PT providing to the subject at least one Borrelia antigen or its fragment.
 XX
 XX Claim 27; SEQ ID NO 64; 121pp; English.

CC The invention describes a method of immunizing a subject comprising
 CC providing to the subject at least one Borrelia antigen or its fragment to
 CC induce an immune response. Also described are: an isolated polynucleotide
 CC comprising a sequence having at least 17 contiguous nucleotides in common
 CC with a sequence not given in the specification; an isolated polypeptide
 CC having at least 5 consecutive amino acids of the sequence not given in
 CC the specification; a vaccine composition comprising at least one Borrelia
 CC antigen or at least one polynucleotide encoding a Borrelia antigen;
 CC screening for at least one test polypeptide or test polynucleotide
 CC encoding a polypeptide for an ability to produce an immune response;
 CC preparing a vaccine; vaccinating a subject; treating a subject infected
 CC with a pathogen; passive immunization; vaccination; detecting Herpesvirus
 CC ; and detecting antibodies to a herpesvirus. The method is useful in
 CC immunizing a subject against Borrelia burgdorferi infection. This is the
 CC amino acids sequence of a Borrelia burgdorferi polypeptide that may be
 CC useful in immunizing a subject against infection by a member of the
 CC Borrelia species.
 XX

SQ Sequence 370 AA;

Query Match 97.2%; Score 1837; DB 8; Length 370;
 Best Local Similarity 97.0%; Pred. No. 1.7e-149; Indels 0; Gaps 0;
 Matches 359; Conservative 6; Mismatches 5;

QY 1 MPQDTISVSLDSRIQASRPNNYPLVYKTAKIKVNDAAANYKILNLTNNYKQIETL 60
 DB 1 MPQDTISVSLDSRIQASRPNNYPLVYKTAKIKVNDAAANYKILNLTNNYKQIETL 60
 QY 61 EKDNNGQDQFGKTKLLKTAMSNFNSSESLKSADLFYKDKPEELKYLKVRHTFV 120
 DB 61 EKDNNGQDQFGKTKLLKTAMSNFNSSESLKSADLFYKDKPEELKYLKVRHTFV 120
 QY 121 VLINTEGNSDGLKIYKDDYDKFTPSIFFVSTKEQEIKELFKDKGNTKERNIAVYS 180
 DB 121 VLINTEGNSDGLKIYKDDYDKFTPSIFFVSTKEQEIKELFKDKGNTKERNIAVYS 180
 QY 181 NKNLNLHLKFIQYLHQASIFHAVNPYGMPLAATPLVDDTVIGKLTAKINFYSLNETG 240
 DB 181 NKNLNLHLKFIQYLHQASIFHAVNPYGMPLAATPLVDDTVIGKLTAKINFYSLNETG 240
 QY 241 LDGVPAPFKEGVDLAGGAIDQFTYHYIKNEAIIELIRIWNKNRQNSKLSALQLSGARDN 300
 DB 241 LDGVPAPFKEGVDLAGGAIDQFTYHYIKNEAIIELIRIWNKNRQNSKLSALQLSGARDN 300
 QY 301 AYTSAEICLLKRFVDRGLIIQYKNLSLTSPTPQLKLELSVNITNFSINAVSLVITTD 360
 DB 301 AYTSAEICLLKRFVDRGLIIQYKNLSLTSPTPQLKLELSVNITNFSINAVSLVITTD 360
 QY 361 IVDYQNSLSA 370
 DB 361 IVDYQNSLSA 370

RESULT 3
 ADU98803
 ID ADU98803 standard; protein; 160 AA.
 XX
 AC ADU98803;
 XX
 XX 24-FEB-2005 (first entry)
 DT
 DE Borrelia burgdorferi antigenic polypeptide seqid 62.
 XX
 XX antibacterial; vaccine; immune stimulation; immunity; antigen;
 KW DNA library.
 XX
 XX Borrelia burgdorferi.
 OS

PN WO2004103269-A2.
XX 02-DEC-2004.
XX 17-OCT-2003; 2003WO-US033056.
XX 18-OCT-2002; 2002US-0419401P.
XX (MACR-) MACROGENICS INC.
XX (TEXA) UNIV TEXAS SYSTEM.
XX Sykes KF, Hale KS, Johnston SA;
XX WPI; 2004-834155/82.
XX N-PSDB; ADU98802.
XX Immunizing a subject against Borrelia burgdorferi infection comprises
PT providing to the subject at least one Borrelia antigen or its fragment.
XX Claim 27; SEQ ID NO 62; 121pp; English.
XX The invention describes a method of immunizing a subject comprising
CC providing to the subject at least one Borrelia antigen or its fragment to
CC induce an immune response. Also described are: an isolated polynucleotide
CC comprising a sequence having at least 17 contiguous nucleotides in common
CC with a sequence not given in the specification; an isolated polypeptide
CC having at least 5 consecutive amino acids of the sequence not given in
CC the specification; a vaccine composition comprising at least one Borrelia
CC antigen or at least one polynucleotide encoding a Borrelia antigen;
CC screening for at least one test polypeptide or test polynucleotide
CC encoding a polypeptide for an ability to produce an immune response;
CC preparing a vaccine; vaccinating a subject; treating a subject infected
CC with a pathogen; passive immunization; vaccination; detecting Herpesvirus
CC ; and detecting antibodies to a herpesvirus. The method is useful in
CC immunizing a subject against Borrelia burgdorferi infection. This is the
CC amino acids sequence of a Borrelia burgdorferi polypeptide that may be
CC useful in immunizing a subject against infection by a member of the
CC Borrelia species.
XX Sequence 160 AA;
SQ
Query Match 41.7%; Score 788; DB 8; Length 160;
Best Local Similarity 98.1%; Pred. No. 1.3e-59;
Matches 157; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 211 LAATPLVDVTVGKLTAKINFVSLNETGLDGVPAFKGVDLGAGIDEQFTYHYKNE 270
DB 1 LAATPLVDVTVGKLTAKINFVSLNETGLDGVPAFKGVDLGAGIDEQFTYHYKNE 60
QY 271 AIIEILIRIWNKNRQNSKLSALQSGARDNAVYTSATIECLLKFVDVDRGLIIQYKNLSLTLS 330
DB 61 AIIEILIRIWNKNRQNSKLSALQSGARDNAVYTSATIECLLKFVDVDRGLIIIEYKNLRLTLS 120
QY 331 PTPQLKLELSVNTITNFNSINAVSLVTTQDIDVYQNSLSA 370
DB 121 PTPQLKLELSVNTITNFNSINAVSLVTTQDIDVYQNSLSA 160
RESULT 4
AA83290
ID AAY83290 standard; protein; 79 AA.
XX AAY83290;
XX 12-SEP-2003 (revised)
XX 16-AUG-2000 (first entry)
XX Borrelia antigenic polypeptide.
XX Diagnosis; detection; immune response; antigen; antibody; immunogen;
XX vaccine; treatment; Lyme borelliosis.
XX Borrelia burgdorferi; ZS7.
XX WO200022134-A2.
XX 20-APR-2000.
XX 12-OCT-1999; 99WO-EF007651.
XX 13-OCT-1998; 98DE-01047142.
XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX Simon M, Wallich R, Kramer M;
XX WPI; 2000-329166/28.
XX N-PSDB; AAZ93798.
XX Diagnosing borelliosis by detecting specific antibodies, using new
PT Borellia antigens expressed preferentially during infection and also
PT useful as immunogens in vaccines.
XX Claim 11; Page 44; 56pp; German.
XX Diagnosis of borelliosis infection can be achieved by detecting anti-
CC Borellia antibodies. The antibodies are detected using a detection
CC reagent that contains many Borellia antigens and reacting this reagent
CC with serum from Borellia infected animals. The antigens have molecular
CC weights of about 9.5, 18, 19, 30, 32, 33, 62, 70, 80, 90, 100 and/or 102
CC kD. The method is particularly used to diagnose Lyme borelliosis. The
CC antigens (or cell lysate or fractions containing them or nucleic acid
CC encoding them) are also useful as immunogens for production of
CC antibodies. They can also be used in vaccines for treatment and
CC prevention of borelliosis. This method of detection is advantageous since
CC it detects immunological structures in Borellia that are expressed
CC selectively in vivo and provides a more reliable diagnosis than existing
CC methods. See GENESEQ records AAZ93790-293807 and AAY83282-183299.
XX (Updated on 12-SEP-2003 to standardise OS field)
XX Sequence 79 AA;
SQ
Query Match 20.8%; Score 392; DB 3; Length 79;
Best Local Similarity 94.9%; Pred. No. 6.7e-26;
Matches 75; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 66 NGQDFGKEKTLTKTAMSNFNSSESLKSADLFYKDKPEELKYLKVRHRTFVLLNT 125
DB 1 NGQDFGKEKTLTKTAMSNFNSSESLKSADLFYKDKPEELKYLKVRHRTFVLLNT 60
QY 126 EGDNSDDGLKIYKDDYDKP 144
DB 61 EGDASDDGLKIYKODYNKP 79
RESULT 5
ADU98759
ID ADU98759 standard; protein; 65 AA.
XX AC
XX ADU98759;
XX 24-FEB-2005 (first entry)
XX Borrelia burgdorferi antigenic polypeptide seqid 18.
XX antibacterial; vaccine; immune stimulation; immunity; antigen;
XX DNA library.
XX Borrelia burgdorferi.
XX WO2004103269-A2.
XX 02-DEC-2004.
XX 17-OCT-2003; 2003WO-US033056.
XX

```

PR 18-OCT-2002; 2002US-0419401P.
XX (MACR-) MACROGENICS INC.
PA (TEXA ) UNIV TEXAS SYSTEM.
XX
PI Sykes KF, Hale KS, Johnston SA;
XX
DR WPI; 2004-834155/82.
DR N-PSDB; ADU98758.
XX
PT Immunizing a subject against Borrelia burgdorferi infection comprises
PT providing to the subject at least one Borrelia antigen or its fragment.
XX
XX Claim 27; SEQ ID NO 18; 121pp; English.
PS
CC The invention describes a method of immunizing a subject comprising
CC providing to the subject at least one Borrelia antigen or its fragment to
CC induce an immune response. Also described are: an isolated polynucleotide
CC comprising a sequence having at least 17 contiguous nucleotides in common
CC with a sequence not given in the specification; an isolated polypeptide
CC having at least 5 consecutive amino acids of the sequence not given in
CC the specification; a vaccine composition comprising at least one Borrelia
CC antigen or at least one polynucleotide encoding a Borrelia antigen;
CC screening for at least one test polypeptide or test polynucleotide
CC encoding a polypeptide for an ability to produce an immune response;
CC preparing a vaccine; vaccinating a subject; treating a subject infected
CC with a pathogen; passive immunization; vaccination; detecting Herpesvirus
CC ; and detecting antibodies to a herpesvirus. The method is useful in
CC immunizing a subject against Borrelia burgdorferi infection. This is the
CC amino acids sequence of a Borrelia burgdorferi polypeptide that may be
CC useful in immunizing a subject against infection by a member of the
CC Borrelia species.
XX
XX Sequence 65 AA;
SQ
Query Match 17.4%; Score 329; DB 8; Length 65;
Best Local Similarity 100.0%; Pred. No. 1.4e-20;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 211 LAATPLVDDTVIGKLTAKINFYSLNLTGDLGVPAFKEGVDLAGGAIDEQFTYHYKNE 270
DB 1 LAATPLVDDTVIGKLTAKINFYSLNLTGDLGVPAFKEGVDLAGGAIDEQFTYHYKNE 60
QY 271 AIEL 275
DB 61 AIEL 65
|||||
RESULT 6
ABU42513
ID ABU42513 standard; protein; 1633 AA.
XX
XX AC ABU42513;
XX
DT 19-JUN-2003 (first entry)
DE Protein encoded by Prokaryotic essential gene #28040.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Staphylococcus epidermidis.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX

```

```

XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyekind JW;
XX FI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA46383.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 70437; 1766pp; English.
PS
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1633 AA;
Query Match 7.3%; Score 137.5; DB 6; Length 1633;
Best Local Similarity 22.6%; Pred. No. 0.043;
Matches 97; Conservative 58; Mismatches 168; Indels 107; Gaps 21;
QY 18 SRPNYNNPLLVYKTAIKVKNK-DAANYKILNLTVNNYKQIETLEKNGNGODQFGKEKT 76
DB 113 SKPRANEAVVTNBSKPKTTEAPTNEBSIAETPKTSTTQQDSTTEKNFSLKDLNLSST 172
QY 77 LLKTAMSNFFNSBSLSKADLFYK-----DKPEELKKYLKVHRHFFVLINTEGDS 130
DB 173 TSKE-----KTDEHSTKQAMSTNKSNDLTNDSTQSEK-----TSSQANNDSTNQ 220
QY 131 D-----DGLKIYKDDYD-----KFKTPSIFVFVSTKEQE-----IKE 162
DB 221 SAPSKQLDSKPSKQVYKTKFNDEPTQDVETTYTKLTPSISTDSSVNDKQDYTSASAS 280
QY 163 LFKDKGNTKERNIAVYNNKNDNLHLKFIS-QYHLQASIFHAV-----NP-YGM--PLAA 213
DB 281 LGVDSNETE-----AITNAVRDNLDLKTASREQINEAIIAELKDKFSDPYGVDTPAL 335
QY 214 TPLVDDTVIGKLTAKINFYSLNLTG-----DGVPAFKEGVDI-----AGGAI----- 258
DB 336 NTSQKTSPhKASGPRMNLMSLAAPNSGKNVDKVTNPTLSLNKNNHANNVWPT 395

```

QY 259 DEQFTY--HYIKNEAI-----IELIRIWNKNNRQNSKLSALQLSGAR 298
 DB 396 NEQFNKANYELDDSIKEGDTFTIKYQYIRPGLELPAI---KTQLRSKDGSIIVANGVY 452
 QY 299 DNATSAIECLLKRPVDRGLIIQYKNL--SUTLSPOLKLELSVNIYFNFINAVSLVI 356
 DB 453 DKT-TNTTYYTFTNYVD-----QYQNTGTFDLIATPKRETAIKDNQNPMEVTIANEW 506
 QY 357 TTQDIVDYON 366
 DB 507 KKOFIVDYGN 516

RESULT 7

AAB18180
 ID AAB18180 standard; protein; 2295 AA.

AC AAB18180;

DT 07-NOV-2000 (first entry)

DE Plasmodium falciparum chromosome 2 related protein SEQ ID NO:37.

KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
 antimalarial; malaria; protozoaside; infection; insecticide.

XX Plasmodium falciparum.

OS WO200025728-A2.

PN 11-MAY-2000.

PF 05-NOV-1999; 99WO-US026796.

PR 05-NOV-1998; 98US-0107131P.

XX (HOFF/) HOFFMAN S.

PA (CARU/) CARUCCI D.

PA (GARD/) GARDNER M.

PA (VENT/) VENTER J C.

XX Hoffman S, Carucci D, Gardner M, Venter JC;

PI WPI; 2000-365347/31.

XX Proteins encoded by chromosome 2 of the human malarial parasite,
 Plasmodium falciparum, useful as antimalarial vaccines and in the
 diagnosis of P.falciparum infection.

XX Disclosure; Page 92-98; 577pp; English.

XX The present invention describes proteins and their fragments (I) encoded
 by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
 CC Also described are: (i) nucleotide sequences (ii) encoding (i); and (2)
 CC vaccines against P. falciparum infection comprising (i) or (ii). (i) and
 CC (ii) are useful for the development of vaccines against P. falciparum
 CC infection. (i) and polyclonal antisera or a monoclonal antibody raised to
 CC immunogens comprising the sequences of (i), are useful in the detection
 CC of infection with P. falciparum. Furthermore, (i) (especially when they
 CC are refined or secreted or membrane proteins) can aid the identification
 CC of drugs to treat or prevent P. falciparum infection, or they can be used
 CC to identify drug resistance in P. falciparum. Sequencing of the
 CC Plasmodium chromosome 2 and the subsequent identification of proteins
 CC encoded by it will help to expand our understanding of parasite biology,
 CC a process hampered by the complexity of the parasitic lifecycle, and
 CC provide new targets for vaccine and drug development. Parasite resistance
 CC to drugs and mosquito resistance to insecticides have led to a resurgence
 CC of malaria in many parts of the world, and there is a pressing need for
 CC vaccines and new drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352
 CC represent nucleotide and protein sequences given in the present
 CC invention, but which are not specifically mentioned within the
 CC specification

XX

SQ Sequence 2295 AA;
 Query Match 7.1%; Score 134.5; DB 3; Length 2295;
 Best Local Similarity 21.5%; Pred. No. 0.13;
 Matches 95; Conservative 57; Mismatches 124; Indels 165; Gaps 23;
 QY 23 YNELLVYTKAKIKVKKDAANYKILNLTNNYKQIE-TLEKDNCGGQDFGKEKTLAKTA 81
 DB 648 YIPLIFNPTKDVFTSINQINIKSINFNTIYELITITIKYKKNFYDD-----LUKCR 701
 QY 82 MSNFFNSSESLKSADLFYKDKPEELKK-----YLVKVRHTFVVLINTEGDSDDGLKIY 137
 DB 702 REIFF-KDRHLENNIM---DKQELKKIIRNLMRIHE-----VSEGNRTNINRY 751
 QY 138 KD-----DYDKFKTPIFFVFSTKEQEIKEI-----FKDKGNTEKERN-----TAVY 179
 DB 752 KKYGYDYDKN--ELYVV-----EKNILNVNDTNTFNFMNKEKDKNVFDINKTMRIVDY 805
 QY 180 SNN-----KDNLH--LK-----FISQYLHQA-----SIFHAVN----- 205
 DB 806 YNNINLNIPTAAIKMKDKIYDQKLKLSNFVEKLKNSICVLSFLYLIGINDNGKLHF 865
 QY 206 PYGMPLAATPLVDDTVIGKLTAKINFYSLNLTGDLGVPAFKGV----- 251
 DB 866 PYGFP-----RNIDFSVKLIREGKGLCNFLSGVLVHINLPIFVNSS 908
 QY 252 -----DLAGGAIDEQFTYHYIKNEAI-----IELIRIWNKNNRQNSKLSAL 292
 DB 909 ISISTEMDDVLEWNDNSINSFFIYYKNNENIRNHDPLSDENRIIPKEDNIKSIISY 968
 QY 293 QLSGARDNAYTSAIECLLKREVDRLIIQYKNLSLTLSPTPOLKLELSVNIY-----NF 347
 DB 969 SLGSKDDFFSKL-----AFTNNVIRLKYKN-----KTNNYTLKDYDFD 1007
 QY 348 SINAVSLVITTDIVDYQNSL 368
 DB 1008 -----TFDKINYKNSV 1018

RESULT 8

AAG82343

ID AAG82343 standard; protein; 1155 AA.

XX AAG82343;

DT 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame protein sequence SEQ ID NO:1780.

XX Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;
 endocarditis.

XX Staphylococcus epidermidis.

XX WO200134809-A2.

PD 17-MAY-2001.

PF 09-NOV-2000; 2000WO-US030782.

PR 09-NOV-1999; 99US-0164258P.

XX (GLAX) GLAXO GROUP LTD.

XX Kimmerly WJ;

XX WPI; 2001-316495/33.

XX N-PSDB; AAB53193.

XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 PT useful for vaccinating against infections, e.g. endocarditis.

XX Claim 18; Page 491-492; 2188pp; English.

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)
 CC and (II) can have antibacterial activity and therefore can be used in
 CC vaccination. The nucleic acids (I) may be used to produce the S.
 CC epidermidis polypeptides (II) via the production of vectors containing
 CC them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH5091 to AAH55098
 CC represent oligonucleotide sequences and primers which are used in the
 CC exemplification of the present invention. N.B. The present invention
 CC specifically claims all the polynucleotide sequences given in the
 CC sequence listing of the present specification, however the sequence
 CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
 CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
 CC for SEQ ID NO:4455 to 4464
 XX
 SQ Sequence 1155 AA;

Query Match 7.0%; Score 132.5; DB 4; Length 1155;
 Best Local Similarity 22.3%; Pred. No. 0.072;
 Matches 96; Conservative 58; Mismatches 169; Indels 107; Gaps 21;
 QY 18 SRPNYNNPLLYVKTAKIKVKNK-DAANYKILNLTVNNYEKQIETLEKDNKGNGQDQFGKBT 76
 DB 113 SKPKANEAVVTNESTKPKTTEAPTVEESIAETPKTSTTQDSTTEKNPDLNLSST 172
 QY 77 LLKTAMSNFFNSSESLSKADLFYK-----DKPEELKKYLKVRHHTFVVLINTEGNS 130
 DB 173 TSKE-----KTDEHSTKQAQMSKNLNDTNDSTQSEK-----TSSQANNDSTNQ 220
 QY 131 D-----DGLKIYKDDYD-----KFTPTSIFFVFSTKEQ-----IKE 162
 DB 221 SAPSKQLDSKPEQVKYTKFNDPTQDVEHTTTKLTPTSISTDSSVNDKQDYSASV 280
 QY 163 LFKDKGNTKERNIAVYNNKNDLHLKPFIS-QYLHQASIFHAV-----NP-YGM--PLAA 213
 DB 281 LGVDSNETE-----AITNAVRDNLDLKAASREQINEAIIAEALKKDFSNPDYGVDTPLAL 335
 QY 214 TPLVDDTVIGLRTAKINPYSILNTEGL-----DGVPAKFGVDI-----AGGAI----- 258
 DB 336 NTQSQNSPKHSGASFPNMLMSLAEPNSGKNVNDKVKITNPTLSLKNKNHANNVWPTS 395
 QY 259 DEQFTY--HYIKNEAI-----IELIRIMNKNRQNSKLSALQLSGAR 298
 DB 396 NEQFNKANYELDDSIKEGDTFTIKYQVIRPGGLEPAI---KTQLRSKDGSIIVANGVY 452
 QY 299 DNAYTSAIECLKRFVDRGLIIQYKNL--SLTSLPTQLKLSLVNITNFSINAVSLVI 356
 DB 453 DKT-TNTTYYTNTYVD-----QYQNTGSPDLIATPKRETAIKDNQNPMEVTIANEVY 506
 QY 357 TTQDVIDYON 366
 DB 507 KKDFIVDIGN 516

RESULT 9
 ID AAY83170
 XX AAY83170 standard; protein; 1802 AA.
 AC AAY83170;
 XX
 DT 24-JUL-2000 (first entry)
 XX
 DE Cell wall protein SdrF.
 XX
 KW SdrF; SdrG; SdrH; coagulase negative; staphylococcus; septicemia;
 KW osteomyelitis; endocarditis; immune response; vaccine; graft; stent;

KW intravenous catheter; heart valve; cardiac.
 XX Staphylococcus sp.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 12 /note= "Position encoded by TAA stop codon"
 FT Misc-difference 28 /note= "Position encoded by TAA stop codon"
 FT Misc-difference 1771 /note= "Position encoded by TAG stop codon"
 FT Misc-difference 1774 /note= "Position encoded by TGA stop codon"
 FT Misc-difference 1793 /note= "Position encoded by TAA stop codon"
 FT Misc-difference 1800 /note= "Position encoded by TGA stop codon"
 XX
 PN WO200012689-A1.
 XX
 PD 09-MAR-2000.
 XX
 PF 31-AUG-1999; 99WO-US019728.
 XX
 PR 31-AUG-1998; 98US-0098443P.
 PR 25-JAN-1999; 99US-0117119P.
 XX
 PA (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
 PA (TEXA) UNIV TEXAS A & M SYSTEM.
 XX
 PI Foster TJ, Hook M, Davis S, Hartford O, McCreia K, Ni Eidhin D;
 XX
 DR WPI; 2000-256637/22.
 DR N-PSDB; AAZ93533.
 XX
 PT Recombinant or synthetic proteins from coagulase-negative staphylococci
 PT useful for prevention, treatment and diagnosis of staphylococcal
 PT infections bind soluble and immobilized fibrinogen.
 XX
 PS Claim 2; Fig 2; 104pp; English.
 XX
 CC Isolated Staphylococcus Sdr cell wall proteins which bind both soluble
 CC and immobilized fibrinogen are useful for treating or preventing
 CC coagulase-negative staphylococcal infection such as septicemia,
 CC osteomyelitis or endocarditis, and for inducing immune responses in
 CC patients. The cell wall proteins are also useful for reducing coagula-
 CC se-negative staphylococci of indwelling medical devices such as
 CC vascular grafts, vascular stents, intravenous catheters, artificial heart
 CC valves and cardiac assist devices. The cell wall associated proteins are
 CC able to inhibit staphylococcal adhesion to immobilised extracellular
 CC matrix or host cells present on the surface of implanted biomaterials
 XX

SQ Sequence 1802 AA;
 Query Match 6.9%; Score 130.5; DB 3; Length 1802;
 Best Local Similarity 22.1%; Pred. No. 0.2;
 Matches 95; Conservative 59; Mismatches 169; Indels 107; Gaps 21;
 QY 18 SRPNYNNPLLYVKTAKIKVKNK-DAANYKILNLTVNNYEKQIETLEKDNKGNGQDQFGKBT 76
 DB 150 SKPKANEAVVTNESTKPKTTEAPTVEESIAETPKTSTTQDSTTEKNPDLNLSST 209
 QY 77 LLKTAMSNFFNSSESLSKADLFYK-----DKPEELKKYLKVRHHTFVVLINTEGNS 130
 DB 210 TSKE-----KTDEHSTKQAQMSKNLNDTNDSTQSEK-----TSSQANNDSTNQ 257
 QY 131 D-----DGLKIYKDDYD-----KFTPTSIFFVFSTKEQ-----IKE 162
 DB 258 SAPSKQLDSKPEQVKYTKFNDPTQDVEHTTTKLTPTSISTDSSVNDKQDYSASV 317
 QY 163 LFKDKGNTKERNIAVYNNKNDLHLKPFIS-QYLHQASIFHAV-----NP-YGM--PLAA 213
 DB 318 LGVDSNETE-----AITNAVRDNLDLKAASREQINEAIIAEALKKDFSNPDYGVDTPLAL 372

QY 214 TPLVDDTVIGKLTAKINFYLLNETGL-----DGVPAFKGVLDL-----AGGAI----- 258
 Db 373 NRSQSKNSPHKSASPRMNLMSLAAPNSGKNVDKVTNPTLSLNKNHANNVIMPTS 432
 QY 259 DEQFTY--HYIKNEAI-----IELIRIWNKNRQNSKLSALQSGAR 298
 Db 433 NEQFNKANYELDDSIKEGDTFTIKYQYIRPGGLELPAI---KTQLRSKDGSIYANGVY 489
 QY 299 DNAYTSAIECLLKRFPVDRGLIIQYKNL--SLTSPTPQLKLELSVNTYFNFSINAVSLVI 356
 Db 490 DKT-TNTTYTFTNVVD-----QYQNTITGSPDLIATPKRETAIKDNQNPMEVTIANEVV 543
 QY 357 TTQDIDVYQN 366
 Db 544 KKDFIVDYGN 553

RESULT 10
 AAY70119
 ID AAY70119 standard; protein; 1802 AA.
 AC AAY70119;
 XX
 DT 06-JUN-2000 (first entry)
 XX
 DE Staph. epidermidis serine-aspartate repeat region protein SdrF.
 XX
 KW Multicomponent vaccine; immunostimulatory; antibacterial; MSCRAMM;
 KW microbial surface components recognising adhesive matrix molecules;
 KW collagen binding protein; CBP; CNA; fibrinogen binding protein;
 KW Clumping factor A; ClfA; Clumping factor B; ClfB; FnBP;
 KW fibronectin binding protein; Staphylococcus infection;
 KW serine-aspartate repeat region protein; SDR protein; SdrF.
 XX
 OS Staphylococcus epidermidis.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 12 /note= "Encoded by in-frame stop codon TAA"
 FT Misc-difference 28 /note= "Encoded by in-frame stop codon TAA"
 FT Misc-difference 1771 /note= "Encoded by in-frame stop codon TAG"
 FT Misc-difference 1774 /note= "Encoded by in-frame stop codon TGA"
 FT Misc-difference 1793 /note= "Encoded by in-frame stop codon TAA"
 FT Misc-difference 1800 /note= "Encoded by in-frame stop codon TGA"
 FT
 PN WO200012131-A1.
 XX
 PD 09-MAR-2000.
 XX
 PF 31-AUG-1999; 99WO-US019727.
 XX
 PR 31-AUG-1998; 98US-0098439P.
 XX
 PA (INHI-) INHIBITEX INC.
 PA (TEXA) UNIV TEXAS A & M SYSTEM.
 PA (QUEB-) QUEEN ELIZABETH COLLEGE DUBLIN.
 XX
 PI Patti JM, Foster TJ, Hook M;
 XX
 DR WPI; 2000-237781/20.
 DR N-PSDB; AA251201.
 XX
 XX Composition used for generating immune response or for inhibiting
 FT microbial colonization in an animal comprises antibodies that bind
 FT collagen binding protein, fibrinogen binding protein and, optionally,
 FT fibronectin binding protein.

PS Claim 8; Fig 3; 115pp; English.

XX The patent discloses multicomponent vaccines containing selected combinations of bacterial binding proteins termed MSCRAMM (microbial surface components recognising adhesive matrix molecules) or their antibodies. A vaccine composition is provided that includes collagen binding protein or peptide, e.g. CNA, a fibrinogen binding protein, preferably Clumping factor A (ClfA) or Clumping factor B (ClfB), and optionally a fibronectin binding protein e.g. FnBP-A. The vaccines are useful for imparting protection against a broad spectrum of Staphylococcal strains and for inhibiting microbial colonisation, especially of Staphylococcus aureus, in an animal. The combinations can also be used to select donor blood pools for the preparation of purified blood products for passive immunisation. The present sequence is a serine-aspartate repeat region protein, SdrF from Staphylococcus epidermidis. The Sdr protein is useful in vaccine preparation in combination with specific bacterial binding proteins. These vaccines can be used to treat a broad spectrum of bacterial infections, including those arising from both coagulase-positive and coagulase-negative bacteria

SQ Sequence 1802 AA;

Query Match 6.9%; Score 130.5; DB 3; Length 1802;
 Best Local Similarity 22.1%; Pred. No. 0.2;
 Matches 95; Conservative 59; Mismatches 169; Indels 107; Gaps 21;

QY 18 SRPNYTPLLVYKTAKIKVVK-DAANYKILNLTNNYEKQIETLEKDNKGQDQFGKEKT 76
 Db 150 SKPKANEAVVTNESTKPKTTEAPTNEESIAETPKSTTQDSTKKNPSUKDNLNSST 209
 QY 77 LLKTAMSNFFNSSSESLSADLFYK-----DKPEELKKYLKVRHRTFVVLINTEGDS 130
 Db 210 TSKE-----KTDEHSTKQAMSTKSNLDTNDSPQSEK-----TSSQANNDSTNQ 257
 QY 131 D-----DGLKIYKDDYD-----KFKTPSIFVFSTKEQE-----IKE 162
 Db 258 SAPSKQLDSKPSQKVKYKTKFNDPTQDVEHTTTKLTSPSVSTDSSVNDKQDYTRSAVAS 317
 QY 163 LFKQKNGTEKERNIAVSNKDNLHLKFIIS-OYLHOASIFHAV-----NP-YGM--PLAA 213
 Db 318 LGVDSNETE-----AITNAVRDNLDLKAASREQINEAIIAEALKKDFSNPDYGVDTPLAL 372
 QY 214 TPLVDDTVIGKLTAKINFYLLNETGL-----DGVPAFKGVLDL-----AGGAI----- 258
 Db 373 NRSQSKNSPHKSASPRMNLMSLAAPNSGKNVDKVTNPTLSLNKNHANNVIMPTS 432
 QY 259 DEQFTY--HYIKNEAI-----IELIRIWNKNRQNSKLSALQSGAR 298
 Db 433 NEQFNKANYELDDSIKEGDTFTIKYQYIRPGGLELPAI---KTQLRSKDGSIYANGVY 489
 QY 299 DNAYTSAIECLLKRFPVDRGLIIQYKNL--SLTSPTPQLKLELSVNTYFNFSINAVSLVI 356
 Db 490 DKT-TNTTYTFTNVVD-----QYQNTITGSPDLIATPKRETAIKDNQNPMEVTIANEVV 543
 QY 357 TTQDIDVYQN 366
 Db 544 KKDFIVDYGN 553

RESULT 11

AAAB18171
 ID AAAB18171 standard; protein; 1979 AA.

AC AAAB18171;

XX 07-NOV-2000 (first entry)

DE Plasmodium falciparum chromosome 2 related protein SEQ ID NO:28.

XX Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;

KW antimalarial; malaria; protozoacide; infection; insecticide.

XX Plasmodium falciparum.


```
CC the invention.
XX
SQ Sequence 575 AA;

Query Match      6.4%; Score 120.5; DB 8; Length 575;
Best Local Similarity 20.6%; Pred. No. 0.29;
Matches 84; Conservative 68; Mismatches 131; Indels 125; Gaps 22;

QY 29 YKTAKIKVNDK-ANYKILNLTNNYKQIETLEKDNQ-GDQPGKKEKTLTKTAMS--- 83
DB 29 YKAWKKEFDROVQKGAQKSMIIAPIVQKRNVEIKDDNGNIVRDNNNGNPKTKPVITGYR 128
QY 84 --NPNSEESLSADLFIYKDKPEELKKYLKHVHTFVVLINTSGDSDGLKLYKDDY 141
DB 129 AHNVEDVAD--TKGKPLITAKD-----LIKTEFENSND-----YKDLI 164
QY 142 DKFK-----TPSI---FFV-----FSTKQEI-----KELPKDK 167
DB 165 NEFKDYINEELTPSVEEKHFLDDPTLSNGAKGYSPKSDIEIVISDDLSYDMRFTLIHEY 224
QY 168 GNTKERNIAVYNNKNDLHLKFISQYLHQASIFHAVNPYGMPLAATPL-----V 217
DB 225 AHSQHLNN--DIGKTQISEHRSLSKEIEAESAYVAVNYGLDTSYSLVSGWGHNIS 282
QY 218 DDTVIGKLTAKINFYSLNET--GLDGVPAKGVLAGGNAIDRQFTHYIKNEAIIEL 275
DB 283 DDELKAHIK----NIHSPAKTTIEINSLPFSQVID---KKLESEM-----NKDVYKD 329
QY 276 TRINKNNKQN--SKLSALQ-----LSGARDNAYTSATCLLKRFPVDRGLIYQKNL 325
DB 330 LSTMTDNLKNGFDKILTIKGNLVNDYGLNISENSYEND-----DPKNIDYKGF 380
QY 326 SLTISPTQLKLEL-----SVNITNFS-----INAVSLVITQD 360
DB 381 N-TNNSQDQAKIELINKHDDSLRDYNTFTQTVNRVNNNTTIFVQDD 427

RESULT 13
ABU26292
ID ABU26292 standard; protein; 598 AA.
XX
AC ABU26292;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #11819.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Campylobacter jejuni.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA30162.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
```

```
PT isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 54216; 1766pp; English.
XX
The invention relates to an isolated nucleic acid comprising any one of
the 6213 antisense sequences given in the specification where expression
of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 598 AA;

Query Match      6.4%; Score 120.5; DB 6; Length 598;
Best Local Similarity 22.5%; Pred. No. 0.3;
Matches 97; Conservative 69; Mismatches 133; Indels 133; Gaps 25;

QY 10 LLSRIQASRPNYNPLLYVTAKIKVNDKAAKYLNTVN---NYEKQIETL----- 60
DB 171 LLD--LKATFEN-LDKADPFKGVNDVNFKEIINNKNISVKNLNNHLENTTHTTTHSMQ 227
QY 61 ---EKDNGNGDQFGKEKTLTKTAMSNFFN--SSESLKSADLFIYKDKPEELKKYLVH 115
DB 228 KTNKDSG-----SLLSQTLKNLDSLSKES-----KHENKVKVS--KIE 267
QY 116 RHTFVLINTSGDSDGLKIYKDDYDKFKTFPIFFVFSTKEQ-----EIKELPKD---K 167
DB 268 EDT-----TDKNTLKNIK--NDEFAKNLTEL--NIKDKNQDNLNKNESKDLNKFNE 318
QY 168 GNTKERNIAVYNNKD--NLHLKFISQYLHQASIFHAVNPYGMPLAATPLVDDT--VIGK 224
DB 319 LNKQKNNLNQENTIQDNKLNKNDQNLNLDKNLN-----KEIVKDTQKLVSN 367
QY 225 LRTAKINFYSLNETGLDGVPAKGVLAGGNAIDRQFTHYIKNEAIIELIRNKK--- 281
DB 368 LTQKDFN----LNKEPKNN--NKENDKIKONFFDQKLNLENLANKTQVVQ-----NKENN 415
QY 282 ----NNRQNSK-----LSALQSLSGARDNAYTSATC 308
DB 416 ANFNNTNNKETTQEQTKTHSENVNDKNSLDELNSLVKDLNKNYTONNARNITPKETLOY 475
QY 309 L---LKRFPVDRGLIYQK---NLSLTLSPTPQLKLELSV-----NITNFSINAVSLVI 356
DB 476 FSQDLKEAVD---QYKAPITKLISITLNPNNLGEVETLIRQGNLHNFNSNANAML 530
QY 357 TQDIVYQNSL 368
DB 531 FIQNAEPKNSL 542
```

XX SQ Sequence 1178 AA;

Query Match 6.3%; Score 119; DB 6; Length 1178;
Best Local Similarity 22.6%; Pred. No. 1.1;
Matches 67; Conservative 50; Mismatches 118; Indels 62; Gaps 16;

QY 14 RIQASRPNY-----NPLVYKTAIKVKNDAANYKILNLT--VNNYKQIETL-- 60
DB 250 KVKQNLAAVYQERDLLETKNQTLKEKRLSRQMDQKQADLEMLTRLSIDYERQIERIHL 309
QY 61 -----EKONGQDOFGKEKTLTKTAMSN---FFNSSESLKSADLFYKDKPBLK 109
DB 310 EYVSKTEKKQSQKSLQSLKTELAKQVLEQLBEEDLKQKQEI-KNVETELS 368
QY 110 KY-----LKVHRHTFVVLINTEGNSDDGLKIYKDDYDKFKTPSFFVFTKEQEIKE 162
DB 369 RFATPDHIIESLREDFVRLMQKADTSNQ-LAVLKAEKMDSRKQES-----ESTAEIKQ 422
QY 163 LFKD--KGNTEKERNIAVYNNKDNLHLKFISQYLHOASIFHAVN-PYG-MPLAATPLVD 218
DB 423 VQADLEKAKDREQRESANFETAKTKVQ-ELLKDVQKTAQLVQNLAAVYAEQEAQMFLLD 481
QY 219 DTVIGKLRITAKI-----NFYSLNETGLDGVPAKGVVD-IAG--GAIDEQTY 264
DB 482 DVKDKKARQSSILKSLKSHSNFYA-----GVRSLVQQAQDKIKGIVGVASEHLTF 530

RESULT 15
ABU35503
ID ABU35503 standard; protein; 1331 AA.
AC ABU35503;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #21030.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Mycoplasma genitalium.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlseen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA39373.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 63427; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: the sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC <http://wipo.int/pub/published/pct/sequences>

RESULT 14
ABU44244
ID ABU44244 standard; protein; 1178 AA.
XX
AC ABU44244;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #29771.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Streptococcus mutans.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlseen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA48114.
XX
DE New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 72168; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: the sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC <http://wipo.int/pub/published/pct/sequences>

nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at icp.wipo.int/pub/published_pct_sequences

Sequence 1331 AA;

Search completed: January 24, 2006, 19:44:54
Job time : 85.6852 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:28:54 ; Search time 14.0368 Seconds
(without alignments)
2536.214 Million cell updates/sec

Title: US-10-688-058-20
Perfect score: 1889
Sequence: 1 MPQDTISVSLDSRIQASRP.....AVSLVITQDIVDYQNSLSA 370

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	803.5	42.5	356	2	F70212	hypothetical prote
2	134.5	7.1	2295	2	B71621	probable membrane
3	128.5	6.8	1979	2	C71622	hypothetical prote
4	127.5	6.7	1034	2	G90591	hypothetical prote
5	127.5	6.7	1127	2	T28317	ORF MSV156 hypothe
6	126	6.7	913	2	T18503	hypothetical prote
7	126	6.7	1726	2	A45948	major merozoite su
8	126	6.7	2166	2	G70163	hypothetical prote
9	125.5	6.6	1070	2	F90106	IAp100 protein [im
10	121	6.4	1701	2	A54498	major merozoite su
11	121	6.4	1726	1	SAZQGM	major merozoite su
12	121	6.4	2269	2	T28677	rhostry protein -
13	120.5	6.4	598	2	D81420	hypothetical prote
14	119.5	6.3	834	2	A97178	probable permease
15	119	6.3	612	2	S61116	probable membrane
16	119	6.3	1206	2	A64207	hypothetical prote
17	119	6.3	1411	2	T18417	hypothetical prote
18	118.5	6.3	834	2	T28250	ORF MSV089 probabl
19	117.5	6.2	3394	2	T18501	hypothetical prote
20	117.5	6.2	5005	2	F82884	hypothetical prote
21	117	6.2	1156	2	B70356	chromosome assembl
22	117	6.2	1272	2	H82926	conserved hypothet
23	117	6.2	1417	2	T18418	hypothetical prote
24	117	6.2	1701	2	A26868	major merozoite su
25	117	6.2	3844	2	T18402	asparagine/asparta
26	116.5	6.2	1191	2	B97116	chromosome segrega
27	116	6.1	1061	2	D98008	conserved hypothet
28	116	6.1	1367	2	T18466	hypothetical prote
29	115.5	6.1	297	2	T28188	hypothetical prote

ALIGNMENTS

RESULT 1

F70212

hypothetical protein BBA46 - Lyme disease spirochete plasmid A/lp54

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004

C:Accession: F70212

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White

son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,

Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:96065943; PMID:9403685

A:Accession: F70212

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-356 <KLE>

A:Cross-references: UNIPROT:O50939; UNIPARC:UPI0000056927; GB:AE000790; NID:g2690224; P1

A:Experimental source: strain B31

C:Genetics:

A:Genome: plasmid

C:Superfamily: Lyme disease spirochete plasmid A/lp54 hypothetical protein BBA46

Query Match 42.5%; Score 803.5; DB 2; Length 356;

Best Local Similarity 46.3%; Pred. No. 4e-45;

Matches 171; Conservative 58; Mismatches 117; Indels 23; Gaps 6;

Qy 1 MPQDTISVSLDSRIQASRPNNYPLVYKTAKIKVWKDAANYKILNLTANNYKQIETL 60

Db |||:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|

1 MPKDTISVSFQNRILAFNNVNYPLLIYKS-----NVLASKHLASVTNLQDLLEKL 53

Qy 61 EKDNGNGQDQFGK-----EKLTKATMSNFFNSSESLKSDLFYIKDPKPELKKYKVH 115

Db |||:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|

54 EMQKAQESDSKKKIAEQLSALKKSSDFF--GQGLKSVDFVYV-NQIKDKDFLKN 110

Qy 116 RHTFVLAINTEGNSDDGLKIYKDDYDKFTPTSPFFVFSTKEQEIKELFKDKGNTKERN 175

Db |||:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|

111 LHPFVWFVNOAED-----VISDFEIRKICNFIVISTKDANLPDFLKGKDKSELKNI 163

Qy 176 IAVYSNNKDNILHLKPISOYLHQASIFHAVNPYGPPLAATPLVDVDTVIGKLTAKINFYSL 235

Db |||:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|

164 IAVYA-GQNHLHLKTAAYLHQASIFHAVNPYGMILNAVVPYVDDSLDSLAKTNINFYSL 222

Qy 236 LNETGLDVPAPKGVLDLACGAIDQFTYHYIKNEAILELIRIWNKNRQNSKLSALOLS 295

Db |||:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|

223 LNETGNDGILAFKGVSLSGDPIDETFFLYIKNEAIKELIRIWNKNRANSKUGSLND 282

Qy 296 GARDNAYTSAIECLIKRFDVDRGLIIQYKNLSITLSPTPQLKLSLVNTYTFNSINAVSLV 355

Db |||:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|

293 GNLPNYTAGLECHFHKLQKGLIVSYKEIKLINSSEGLSLSLVALKYNDSSFNSVNI 342

Qy 356 ITTQDIVDY 364

Db 343 ITTQBEINEY 351
||||:|:|

RESULT 2

B71621
probable membrane associated protein PFB0190c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004
C:Accession: B71621
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
Perte, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743; PMID:9804551
A:Accession: B71621
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2295 <GAR>
A:CROSS-references: UNIPROT:O9TY98; UNIPARC:UPI000007C03B; GB:AE001379; GB:AE001362; NID
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB0190c

Query Match 7.1%; Score 134.5; DB 2; Length 2295;
Best Local Similarity 21.5%; Pred. No. 1.8;
Matches 95; Conservative 57; Mismatches 124; Indels 165; Gaps 23;

QY 23 YNPLLVYKTAIKVNDKDAANYKILNLTNNYKQIE-TLEKNGNGQDQFGKEKYLKTA 81
Db 648 YIELFNPFTKOVFTSQINIKSKINFFNIYELITITIKENKNFYDD-----LLKCR 701
QY 82 MSNFFNSSSESLKADLFYKDKPELKK---YLVKVRHTFVVLINTEGNSDDGLKTY 137
Db 702 REIFF-KORHLENNIM---DQBELKKNIRNLMRIHE-----VSNEGNNRNTINRKY 751
QY 138 KD-----DYDKFTPSIFFVFSTKEQIKEL-----FKDKGNTKEKERN-----IAVY 179
Db 752 KKGTYDYDKMN--ELYVV-----EKNILNVNDTNTFNFNKNKEKDKNYFDINKTWRIYDY 805
QY 180 SNN-----KQNLH--LK-----FISQVYLHQA-----SIFHAVN-----205
Db 806 YNNINLNIPTPAIRKMDKIYDQLKLLRSNFVEKLNKESICVLSFLYLTIGINDNGKLFH 865
QY 206 PYGMPLAATPLVDVTIGKLTAKNFYSLNLTGLDGVPAKEGV-----251
Db 866 PYGFP-----RNIDFSVKLIREGDKGLCNFLSGVLVHINLPIFVNSS 908
QY 252 -----DLAGGAIDEQFTYHYIKNEAI-----IELIRIWNKNRQNSKLSAL 292
Db 909 ISISTEMNDVLEMDNSINSFYIYKKNENIRNHDFLSDENRIIPRKEDNIKSKIISY 968
QY 293 QLSGARDNAYTSAECLLKRFDVRGLIIQYKNLSUTLSTPQKLELSVNITY-----NF 347
Db 969 SLGSKDDPFSLK-----AFTNNVIRLKYKN-----KTNNTYLYKDYDFD 1007
QY 348 SINAVSLVITTDIDVDYQNSL 368
Db 1008 -----TFDKINYNKSV 1018

RESULT 3

C71622
hypothetical protein PFB0145c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004
C:Accession: C71622
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
Perte, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743; PMID:9804551
A:Accession: C71622

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1979 <GAR>
A:CROSS-references: UNIPROT:O96133; UNIPARC:UPI0000007EA68; GB:AE001375; GB:AE001362; NID
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB0145c

Query Match 6.8%; Score 128.5; DB 2; Length 1979;
Best Local Similarity 20.9%; Pred. No. 3.6;
Matches 77; Conservative 57; Mismatches 128; Indels 107; Gaps 16;

QY 3 QDTISVSLDSRIQASR-----PNYINPLL-----VYKTAK--IKVKNDAAN-- 42
Db 242 RETISIEIIDIKKHLEKLIKIEIKKEDLENLNKKGLSKENVLKEKGVCKEKNETINSL 301
QY 43 -----YKILNLTNNYKQIE-TLEKNGNGQDQFGKEKYLKLTAMSNFFNSSSES 92
Db 302 NDNIIEKEKKYKLLYELEENKQITDLINKOEKEKEKEKEKEK-----EREKEK 352
QY 93 LKSADLFYKDKPELKKYLKVH-----RHTFVVLINTEG-----NSDDGL 134
Db 353 EKEYDTLIKELDKESISILEKVIKVRMDIEKREHNFLEHEDQLKDKNSFVKNQOL 412
QY 135 KIYKDDYDKFTPSIFFVFSTKEQIKEL-----FKDKGN-----TEKERNIAVYS--NNKD 184
Db 413 KVKYCEIKNLKTE-----LEKKEKELKDIEVNSKEEINKLINQLNEKEKEKQILAFKNHKE 467
QY 185 NLH-LKFIQSQVYLHQAIFHAVNPYGMPLAATPLVDVTIGKLTAKNFYSLNLTGLD 243
Db 468 BIHGLK-----BELKESVKITKITQELQEMVDIKQKELDQ 503
QY 244 VPAKEGVDLAGGALD---EQFTYHYIKNEAI--IELIRIWNKNRQNSKLSALQLSGARDN 300
Db 504 LQE-KYNAQIESISIELSKKEKEYNQYKNTYIEINLNLEKLEETNKETYNLQ-----N 556
QY 301 AYTSAEICL 309
Db 557 NYTWEINML 565

RESULT 4

G90591
hypothetical protein MYPU 6390 [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C:Species: Mycoplasma pulmonis
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: G90591
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A:Reference number: A99512; MUID:21267165; PMID:11353084
A:Accession: G90591
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1034 <KUR>
A:CROSS-references: UNIPROT:Q98PT0; UNIPARC:UPI000000D4633; GB:AL445566; PID:gl40900054; P
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPU 6390
A:Genetic code: SGC3

Query Match 6.7%; Score 127.5; DB 2; Length 1034;
Best Local Similarity 23.8%; Pred. No. 1.8;
Matches 94; Conservative 67; Mismatches 141; Indels 93; Gaps 21;

QY 32 AKIKVNDKDAANYKIL-----NLTNNYKQIE-TLEKNGNGQDQFGKEKYLKLTAMS--- 83
Db 227 AQVKISKKEIKNKVLEDKNNLKI-KYDKRINKAEYSYKDYFEQNKQIKESLKHANEKIA 285
QY 84 -----NFFNSE-BLSKS-----ADLFYKDKPEELKKYLKVH-RHTFVVLINT 125
Db 286 QNKDKINELNKNKLEKIKALAEKRALKNYLYFKGLILAKKYKYLKYLINHQKVLICN 345

QY 126 --EGDNS--DGLKLYKDDYDKFTPTSPFFVFTSTKEQIK-----ELPK-----DKGNT 171
 Db 346 YYEKDKTQIDQLWKYKYN-----FLNKVDDQVQNDYFELYKKIINVVENDNSY 395
 QY 172 KERNIAVYNNKNDLHLKAFISOLYHOASIFRAVNPYGMPLAATPLVDVTVIGKLTAKIN 231
 Db 396 KQVVKKIINKSYNFIKY---ELKKNALFHLKSHQWQNL-----LIKQSSYSGD 444
 QY 232 FYSL-----LNETGLDVPAPFEGVDLAGAIDQFTYHYIKNEAIIELIRWKNRQ-- 285
 Db 445 FVEVKSALSOVIDSYNEINNSIIEKQKILEELFNQYAKN-----NLENKQVF 494
 QY 286 NKSLSALQISGARDNAVYSAIECLKR-----FVDRGLIIOYKML--SLTI-SPTP 333
 Db 495 QSKVNLKVDYLLLEKA---KVKKLAKKEITKALVFTNKKLKIDFKKVNLSKLENPKY 551
 QY 334 QLKLELSVNIITVFNINAVSLVITQDIDVDYQNSL 368
 Db 552 QNKLLKTNLGRFLSKKLLHKIYQSKILEAQSI 586

RESULT 5
 T28317
 ORF MSV156 hypothetical protein - Melanoplus sanguinipes entomopoxvirus
 C:Species: Melanoplus sanguinipes entomopoxvirus
 C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
 C:Accession: T28317
 R:Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.
 J. Virol. 73, 533-552, 1999
 A:Title: The genome of Melanoplus sanguinipes entomopoxvirus.
 A:Reference number: 220484; MUID:99102612; PMID:9847359
 A:Accession: T28317
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1127 <AFO>
 A:Cross-references: UNIPROT:Q9YVT6; UNIPARC:UPI00000F6900; EMBL:AF063866; NID:g4049647;
 C:Genetics:
 A:Note: MSV156

Query Match 6.7%; Score 127.5; DB 2; Length 1127;
 Best Local Similarity 21.1%; Pred. No. 2;
 Matches 96; Conservative 68; Mismatches 142; Indels 149; Gaps 21;

QY 32 AKIKVNDKAAANYKILNLT---VNNYKQIETLEKNGQDQFGKERTLLKTAMNFFNS 88
 Db 662 ASITNNIENISNKIKDLNEFIISNDSKELDEIRKYQ-QFDKIKAMNTEVKSFNT 720
 QY 89 SEESLSAD-----LFYDKPEELKYLKVHRTFVVLINTEG-----NSDD 132
 Db 721 LQKDIDSISKNINELTNAVYDIINTKANDLDDKLNNGVSEFNKLYNNSADLLDTIQKNDE 780
 QY 133 GLIKYKDDYDKPTPSI-----FFVFTKE-----QEIKELPK 165
 Db 781 KVKQLNEVLEKKNQKQSIENDIVNNFIKELIKFNNTNKSLSNELLTNDINDIKIFLYK 840
 QY 166 DKGNTKERNIAVYNNKNDL-----HLKFIISOYL-----HQASIFHAVNPYGMPLA 212
 Db 841 ELNKISTNNLLKIYKNEIDVNNKLSIVIENTQFINSFISIEFNOGTSITSHINFLNTLA 900
 QY 213 APPLVDDTVIGKL-----RTAKINFYSLNLTGLDGVPA 246
 Db 901 G---IND-VINKNLKIMADTTRGTGTNRIDRBNKQISSENKSKQFNKNEKDKLKLIS 956
 QY 247 FKEGV---DLAGG-----AIDEQFTYHYIKNEAIIELIRWKNRQN 286
 Db 957 FNDKLNKYNISAGYTEYNNIEHECLKYLYAVSDQ-EYRYK--FLIHNKIDR-NKN 1012
 QY 287 SKLSALQISGARDNAVYSAIECLKRFDVDRGLIIOYKNSLTLSPQPK----- 336
 Db 1013 SMLQFLKL-----FLKRYI---IYDLKILSDTDI-NIQIKQYLYNKLKN 1054
 QY 337 -----LELSVNIITVFNINAVSLV-ITTDIVDY 364

Db 1055 EYKLTPEQFINQTYGFTMPIFNILNISTNMVFDY 1089

RESULT 6
 T18503
 hypothetical protein C0750w - malaria parasite (Plasmodium falciparum)
 C:Species: Plasmodium falciparum
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T18503
 R:Lawson, D.; Bowman, S.; Barrell, B.
 submitted to the EMBL Data Library, August 1997
 A:Reference number: Z18935
 A:Accession: T18503
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-913 <LAW>
 A:Cross-references: UNIPROT:O77386; UNIPARC:UPI000017CC3C; EMBL:Z98551; NID:el331903; PI
 C:Genetics:
 A:Map position: 3
 A:Introns: 384/3
 A:Note: C0750w

Query Match 6.7%; Score 126; DB 2; Length 913;
 Best Local Similarity 23.1%; Pred. No. 1.9; 130; Indels 64; Gaps 15;
 Matches 74; Conservative 52; Mismatches 52; Indels 64; Gaps 15;

QY 8 VSLDSRIQASRPVNPYLLVYKTAK-----IKVNDKAAANYKILNLTVNNYKQIETLEK 62
 Db 510 VSLD---VLSRSNAHG--LIYIDGKNNTSINLNNKNNNNNNNNNNNNNNNNNIS 564
 QY 63 DNGNGQDQFGKERTLLKTAMNFFNSSELSKADLFYKDKPBELK-KYLKVHRTFV 121
 Db 565 NNSNNSNFGHEKVI-TIIPYGPISPTIKIE--FLNKQNNNEKWAYLRLOKNEY-I 620
 QY 122 LINTEGNSDD---GLKIYKDDYDKPK-----TPSIFVSTKEQEIKELPKDKGN 169
 Db 621 MKKVGINSSDIFLSYKYKYKQEKIKFILNDFSETTYLGFIQCTKBEIKKSLCTISLV 680
 QY 170 TEKERNIAVYNNKNDLHLKFIISOYLHQASIFHAVNPYGMPLAATPLVDVTVIGKLTAK 229
 Db 681 DSHKNNMAINLFTNNH---ILFLYDQSPRKIQPY-----NQIISFRFVL 726
 QY 230 I-NFYSLNLTGLDGVPA-----FKEGVDLAGAIDQFTYHYIKNEAIIELI 276
 Db 727 FKNFHLFQENKINLYLLKCMSSSSSYNTQKNIELSYLTNEQVSHDFHTPGI--LI 784
 QY 277 R-----IWNKNRQNSK 288
 Db 785 RPRTNISHNDYQKNKGNHK 804

RESULT 7
 A45948
 major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum)
 C:Species: Plasmodium falciparum
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 31-Dec-2004
 C:Accession: A45948
 R:Chang, S.P.; Kramer, K.J.; Yamaga, K.M.; Kato, A.; Case, S.E.; Siddiqui, W.A.
 Exp. Parasitol. 67, 1-11, 1998
 A:Title: Plasmodium falciparum: Gene structure and hydropathy profile of the major merozo
 A:Reference number: A45948; MUID:89005525; PMID:3049134
 A:Accession: A45948
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1726 <CHA>
 A:Cross-references: UNIPROT:Q25922; UNIPARC:UPI0000177F89; GB:M37213
 C:Superfamily: G surface protein
 C:Keywords: surface antigen

Query Match 6.7%; Score 126; DB 2; Length 1726;
 Best Local Similarity 21.9%; Pred. No. 4.3;
 Matches 94; Conservative 64; Mismatches 145; Indels 126; Gaps 22;

A;Status: preliminary
A;Molecule type: DNA

A;Residues: 2131-2269 <KE2>
A;Cross-references: UNIPARC:UPI000017B646; GB:M34283

Query Match 6.4%; Score 121; DB 2; Length 2269;
Best Local Similarity 19.8%; Pred. No. 13;
Matches 78; Conservative 69; Mismatches 138; Indels 108; Gaps 17;

QY 29 YKTIKIKVKK-----DAANYKILNLTNVNVE-----KQETLE-----K 62
DB 1114 FTKIILKINEIRTKSDCLKEINDLEKQISNLSDTQETKLTENGKQLKTLBELLESKK 1173
QY 63 DNGNGODQFGKEKTLTKTAMSNFFNSSEESLSADLFYI-----KQPEELKKYL 112
DB 1174 OKKNIEDQ-KKELDEVNSKIKNIENVQHKKNGYIGIVEKINEIATAKTKNQIESTKELI 1232
QY 113 KVHRHFFVLINT-----EGNSDDGLKIYKOD-----YDKF-KTPSIF--FVFSTKEQEI-- 160
DB 1233 KPTIOHIISSFNANDLEGIDSDENLGKYNTEMGNTYEEFIKSYNLITNYLETVSKESITY 1292
QY 161 KFLFKDKGNTKERNIAVYNNKDNHLKLFISQYLHQASIFHAVNPYGMPLAATPLVDVT 220
DB 1293 NOIQNKRIDTQKELLKNENVNKAQSYLDYIKENEFDRIVTHFKK----- 1337
QY 221 VIGKLTAKINF---YSLINETGLDGPAPKEGVDLAGGADEQ----- 261
DB 1338 ---KLNTVNDNPKNEYSKVE-GFDNI---SNSINTVKNSTDSLLNLTQTKEMYANI 1390
QY 262 -----FTVHYIKVEAIELIRIWNKNNRQNSKLSALQLSGARDNAYTSAIECLLKRFVDR 316
DB 1391 VNNYTSYKYEAENFRNIPKLANTLNTIKNSSGIDLS-----KDIKIALISYLD- 1441
QY 317 GLIIQVKNLSLSTLPTPOLKLSLVNITVNFSI 349
DB 1442 ----SKTEDTLIFIPSPQKTTYTKISDSYSI 1470

RESULT 13
D81420
hypothetical protein Cj0041 [imported] - Campylobacter jejuni (strain NCTC 11168)
C;Species: Campylobacter jejuni
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: D81420
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrel
Nature 403, 665-668, 2000
A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A;Reference number: A81250; MUID:20150912; PMID:10688204
A;Accession: D81420
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-598 <PAR>
A;Cross-references: UNIPROT:Q9PU70; UNIPARC:UPI00000C2013; GB:AL139074; GB:AL111168; NID
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: Cj0041

Query Match 6.4%; Score 120.5; DB 2; Length 598;
Best Local Similarity 22.5%; Pred. No. 2.4;
Matches 97; Conservative 69; Mismatches 133; Indels 133; Gaps 25;

QY 10 LLDSDRIQSRPNYNNPLLVYTKAKIKVNDKDAANYKILNLTNVN---NYEKQIETL----- 60
DB 171 LLD--LKATFPN-LDKADFFKAVDVFKEIINNKSIVSKNLENLNTHTTTHSHMQ 227
QY 61 ----EKDNGNGODQFGKEKTLTKTAMSNFFN--SSEESLSADLFYIYKQPEELKKYLKVH 115
DB 228 KTNKDSG-----SLLSQTLKNLDSILSSKES-----KHEKNDKVK-S-XIE 267
QY 116 RHTFVVLNTEGNSDDGLKIYKDDYDKETPSIFVFSTKEQ-----EIKELFKD--K 167
DB 268 EDT-----TDAKNTLNKIK--NDEFAKVLTEL--NIKKNQDNLNKESKDLNKNKE 318
QY 168 GNTKERNIAVYNNKND-NLHLKFISQYLHQASIFHAVNPYGMPLAATPLVDVT--VIGK 224

DB 319 LNKQEKNNLNQENIQDNKLNKNDQNLNLDKNLN-----KEIVKDTQKLVSN 367
QY 225 LRTAKINPYSLLNETGLDGPAPKEGVDLAGADEQFTYHYIKVEAIELIRIWNK--- 281
DB 368 LTQKDFN---LNKEPKNN---NKENDKIQKQFFDQKLNFNENLNKTVVQ-----NKENN 415
QY 282 ----NNRNSK-----LSALQLSGARDNAYTSAIEC 308
DB 416 ANFNNNNTNNKETFTQEGTKTHSENVKNSLDELNSLVKDLNKNVTONNARITPKETIQY 475
QY 309 L---LKRFPVDRGLIIQYK---NLSLTSLPTPOLKLELSV-----NITYNFINSINAVSLVI 356
DB 476 FSQDLKEAVD---QYKAPITKLSITLNPNNLGEVETLTIQRGNNLHNFNSNANAVNL 530
QY 357 TTQIDIVDYQNSI 368
DB 531 FIONQAEPKNSL 542

RESULT 14
A97178
probable permease [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: A97178
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: A97178
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-834 <KUR>
A;Cross-references: UNIPROT:Q97GW0; UNIPARC:UPI00000CA47D; GB:AE001437; PIDN:AAK80212.1,
A;Experimental source: Clostridium acetobutylicum ATCC924
C;Genetics:
A;Gene: CAC2255

Query Match 6.3%; Score 119.5; DB 2; Length 834;
Best Local Similarity 23.5%; Pred. No. 4.4;
Matches 92; Conservative 62; Mismatches 137; Indels 101; Gaps 22;

QY 6 ISVSLDSRIQASRPYNNPL--VYKTAKIKVND--AANYKIL---NLTVNNEYKQIE 58
DB 480 ISISLMTIMYMTNSLKDITLTYEYKLNLPDVSIDINSQNVVLLKCSRIINNYDKKTN 539
QY 59 TLEKD--NGNGODQFGKEKTLTKTAMSNFFNSSEESLSADLF-TYKDKPEELKKYLKVH 115
DB 540 VIRKYYINTNLNDDFSKKIDLLCVDPPNF-----KSYDNYLVYTNKQKLDQL--- 587
QY 116 RHTFVVLNTEGNSDDGLKIYKOD---YDKETPSIFVFSTKEQ-----IKELFK 165
DB 588 ----NNNEDGIIISEKNSRRYNNLKGDTISLYSSNSEKEDFKILSVVNAGPM 635
QY 166 DKGNT-----KERNIAVYNNKDNHLKLFISQYLHQASI-FHAVNPY-GMPLAATP 215
DB 636 NMGNINLISFKAALKHFVNR-YSNE-----FFISTSTNTSSIKKOLIKKFKGFPVSVST 688
QY 216 LVDDTVIGKLTAKINFYSLNLTGLDGPAPKEGVDLAGGAIDQFTYHYIKVEAIEL 275
DB 689 KQDN--VNKDAASSKEFISLFN-----IF-----SFTSMIALIGILSN 725
QY 276 IRI--WNKNNRQNSKLSALQLSGARDNAYTSAIECLLKRFVDRGLIIQYKNL-----SLTL 329
DB 736 ISISFVQRKEIASLISIGLSSTGKN-----LMILFESFTE-GLIAFFITFLSSITWTL 779
QY 330 SPTPOLKLSLVNITYNFINSIN-----AVSLVI 356
DB 780 LSDIPKYLVLNITNITYPFSPFWAALSVSIIIM 811

Search completed: January 24, 2006, 19:59:16
Job time : 15.0368 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:27:53 ; Search time 79.505 Seconds
(without alignments)
3283.383 Million cell updates/sec

Title: US-10-688-058-20
Perfect score: 1889
Sequence: 1 MPQDTISVSLDSRIQASRP.....AVSLVITQDIVYQNSLSA 370

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1889	100.0	370	Q9S096 BORBU	Q9S096 borrelia bu
2	1884	99.7	370	Q5XYW3 BORGA	Q5XYW3 borrelia ga
3	1878	99.4	370	Q9S0C8 BORBU	Q9S0C8 borrelia bu
4	1837	97.2	370	Q9R2W7 BORBU	Q9R2W7 borrelia bu
5	1829	96.8	370	Q9S003 BORBU	Q9S003 borrelia bu
6	1828	96.8	370	Q9S0F9 BORBU	Q9S0F9 borrelia bu
7	1755	92.9	362	Q9S051 BORBU	Q9S051 borrelia bu
8	1582.5	83.8	373	Q5XZ68 BORGA	Q5XZ68 borrelia ga
9	803.5	42.5	356	Q5O939 BORBU	Q5O939 borrelia bu
10	421	22.3	106	Q5XYV3 BORGA	Q5XYV3 borrelia ga
11	382.5	20.2	156	Q6AS11 BORGA	Q6AS11 borrelia ga
12	140.5	7.4	76	Q6AS12 BORGA	Q6AS12 borrelia ga
13	140	7.4	2670	Q7YZ02 CRYPV	Q7YZ02 cryptospori
14	140	7.4	2711	Q5CWH1 CRYPV	Q5CWH1 cryptospori
15	137	7.3	1057	Q7RB65 PLAYO	Q7RB65 plasmodium
16	136	7.2	1158	Q7RC66 PLAYO	Q7RC66 plasmodium
17	136	7.2	1356	Q5ZZZ2 MYCHY	Q5ZZZ2 mycoplasma
18	135.5	7.2	654	Q9RLP9 MYCHO	Q9RLP9 mycoplasma
19	134.5	7.1	1347	Q4Z0S9 PLABE	Q4Z0S9 plasmodium
20	134.5	7.1	2295	Q9TY98 PLAF7	Q9TY98 plasmodium
21	134	7.1	872	Q5TIA9 HUMAN	Q5TIA9 homo sapien
22	134	7.1	875	Q8NDG7 HUMAN	Q8NDG7 homo sapien
23	134	7.1	1006	Q4YZ95 PLABE	Q4YZ95 plasmodium
24	134	7.1	1012	Q5TIB0 HUMAN	Q5TIB0 homo sapien
25	134	7.1	1309	Q8TJEB PLAF7	Q8TJEB plasmodium
26	133.5	7.1	1093	Q5ANT5 D1C1I	Q5ANT5 dictyosteli
27	133.5	7.1	1262	Q4YSN7 PLABE	Q4YSN7 plasmodium
28	132.5	7.0	1464	Q8TIF6 PLAF7	Q8TIF6 plasmodium
29	132.5	7.0	1633	Q8CWP4 STAEF	Q8CWP4 staphylococ
30	131.5	7.0	984	Q86IF8 D1C1I	Q86IF8 dictyosteli
31	131.5	7.0	996	Q559A5 D1C1I	Q559A5 dictyosteli

32 131.5 7.0 1186 2 Q895M7 CLOTE
33 131 867 1 IF2 BUCAP
34 130.5 6.9 798 2 Q7RQ74 PLAYO
35 130.5 6.9 1161 2 Q54NR5 D1C1I
36 130.5 6.9 1733 2 Q9K114 STAEF
37 130.5 6.9 2226 2 Q97225 PLAF7
38 129.5 6.9 943 2 Q8ILY0 PLAF7
39 129.5 6.9 1877 2 Q9XXW1 9APIC
40 129.5 6.9 1879 2 Q7PDP1 PLAYO
41 129.5 6.9 1936 2 Q8IDT8 PLAF7
42 129 6.8 456 2 Q4YRA3 PLABE
43 129 6.8 565 2 Q4U973 THEAN
44 129 6.8 1046 2 Q4YRU4 PLABE
45 128.5 6.8 373 2 Q630X2 BACCZ

ALIGNMENTS

RESULT 1

Q9S096 BORBU
ID Q9S096 BORBU PRELIMINARY; PRT; 370 AA.
AC Q9S096;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein BBO11.
GN OrderedLocustNames=BBO11;
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid cp32-7.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OK NCBI_taxID=139;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=20138354; PubMed=10672174;
RA Casjens S., Palmer N., van Vugt R., Huang W.M., Stevenson B., Rosa P.,
RA Lathigra R., Sutton G.G., Peterson J.D., Dodson R.J., Haft D.H.,
RA Hickey E.K., Gwinn M.L., White O., Fraser C.M.;
RT "A bacterial genome in flux: the twelve linear and nine circular
RT extrachromosomal DNAs in an infectious isolate of the Lyme disease
RT spirochete Borrelia burgdorferi.";
RI Mol. Microbiol. 35:490-516(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A.,
RA Artiach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RI Nature 390:580-586(1997).
RL Nature 390:580-586(1997).
DR EMBL; AE001579; AAF07607.1; -; Genomic_DNA.
DR TIGR; BBO11; -;
DR InterPro; IPR008505; DUF787.
DR Pfam; PF05619; DUF787; 1.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 370 AA; 42103 MW; 62FB9015CFEBCD6 CRC64;

Query Match 100.0%; Score 1889; DB 2; Length 370;
Best Local Similarity 100.0%; Pred. No. 4.2e-111;
Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPQDTISVSLDSRIQASRPYNNPLLYVTAKIKVNDKDAANYKILNLTNNYKQIETL 60
DB 1 MPQDTISVSLDSRIQASRPYNNPLLYVTAKIKVNDKDAANYKILNLTNNYKQIETL 60

```
QY 61 EKDNNGQDQFGKEKTLTKTAMSNFNSSESLKSADLFYKDKPEELKKYLKVHRHTFV 120
DB 61 EKDNNGQDQFGKEKTLTKTAMSNFNSSESLKSADLFYKDKPEELKKYLKVHRHTFV 120
QY 121 VLINTEGNSDGLKIYKDDYDKFKTPSPISFFVSTKEQBIKELFKDKGNTKERNIAVYS 180
DB 121 VLINTEGNSDGLKIYKDDYDKFKTPSPISFFVSTKEQBIKELFKDKGNTKERNIAVYS 180
QY 181 NNKDNHLKFIISOYLHQASIFHAVNPGMPLAATPLVDDTVIGKLTAKINFYSLNETG 240
DB 181 NNKDNHLKFIISOYLHQASIFHAVNPGMPLAATPLVDDTVIGKLTAKINFYSLNETG 240
QY 241 LDGVPAPFKEGVDLAGGAIDQFTYHYIKNEAIIELIRIWNKNRQNSKLSALQLSGARDN 300
DB 241 LDGVPAPFKEGVDLAGGAIDQFTYHYIKNEAIIELIRIWNKNRQNSKLSALQLSGARDN 300
QY 301 AYTSAEICLLKRFVDRGLIIQYKNLSLTSPQPKLELSVNITNFNSINAVSLVITQD 360
DB 301 AYTSAEICLLKRFVDRGLIIQYKNLSLTSPQPKLELSVNITNFNSINAVSLVITQD 360
QY 361 IVDYQNSLSA 370
DB 361 IVDYQNSLSA 370
RESULT 2
Q5XW3_BORGA
ID Q5XW3_BORGA PRELIMINARY; PRT; 370 AA.
AC Q5XW3;
DT 25-OCT-2004 (TremBLrel. 28, Created)
DT 25-OCT-2004 (TremBLrel. 28, Last sequence update)
DE Hypothetical protein.
GN ORFNames=SGP168;
OS Borrelia garinii PBI.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=290434;
RN [1]_TaxID=290434;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBI;
RA Glockner G., Lehmann R., Romualdi A., Pradella S.,
RA Schulte-Spechtel U., Schilhabel M., Wilske B., Suhnel J., Platzer M.;
RT "Comparative analysis of the Borrelia garinii genome.";
RL Nucleic Acids Res. 32:6038-6046(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBI;
RA Gloeckner G., Schilhabel M., Lehmann R., Platzer M.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY722923; RAU86019.1; -; Genomic_DNA.
DR InterPro; IPR008505; DUF787.
DR Pfam; PF05619; DUF787; 1.
KW Hypothetical protein.
SQ SEQUENCE 370 AA; 42173 MW; 62FB9015CEBFC9C3 CRC64;
Query Match 99.7%; Score 1884; DB 2; Length 370;
Best Local Similarity 99.7%; Pred. No. 8.7e-111;
Matches 369; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MPQDTISVSLDSRIQASRPNNYPLVYTKAKIKVNDAAANYKILNLTNNYKQIETL 60
DB 1 MPQDTISVSLDSRIQASRPNNYPLVYTKAKIKVNDAAANYKILNLTNNYKQIETL 60
QY 61 EKDNNGQDQFGKEKTLTKTAMSNFNSSESLKSADLFYKDKPEELKKYLKVHRHTFV 120
DB 61 EKDNNGQDQFGKEKTLTKTAMSNFNSSESLKSADLFYKDKPEELKKYLKVHRHTFV 120
QY 121 VLINTEGNSDGLKIYKDDYDKFKTPSPISFFVSTKEQBIKELFKDKGNTKERNIAVYS 180
DB 121 VLINTEGNSDGLKIYKDDYDKFKTPSPISFFVSTKEQBIKELFKDKGNTKERNIAVYS 180
```

```
QY 181 NNKDNHLKFIISOYLHQASIFHAVNPGMPLAATPLVDDTVIGKLTAKINFYSLNETG 240
DB 181 NNKDNHLKFIISOYLHQASIFHAVNPGMPLAATPLVDDTVIGKLTAKINFYSLNETG 240
QY 241 LDGVPAPFKEGVDLAGGAIDQFTYHYIKNEAIIELIRIWNKNRQNSKLSALQLSGARDN 300
DB 241 LDGVPAPFKEGVDLAGGAIDQFTYHYIKNEAIIELIRIWNKNRQNSKLSALQLSGARDN 300
QY 301 AYTSAEICLLKRFVDRGLIIQYKNLSLTSPQPKLELSVNITNFNSINAVSLVITQD 360
DB 301 AYTSAEICLLKRFVDRGLIIQYKNLSLTSPQPKLELSVNITNFNSINAVSLVITQD 360
QY 361 IVDYQNSLSA 370
DB 361 IVDYQNSLSA 370
RESULT 3
Q9S0C8_BORBU
ID Q9S0C8_BORBU PRELIMINARY; PRT; 370 AA.
AC Q9S0C8;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE Hypothetical protein BBM11.
GN OrderedLocNames=BBM11;
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=139;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=20138354; PubMed=10672174;
RA Casjens S., Palmer N., van Vugt R., Huang W.M., Stevenson B., Rosa P.,
RA Lathigra R., Sutton G.G., Peterson J.D., Dodson R.J., Hickey E.K.,
RA Hickey E.K., Gwinn M.L., White O., Fraser C.M.;
RT "A bacterial genome in flux: the twelve linear and nine circular
RT extrachromosomal DNAs in an infectious isolate of the Lyme disease
RT spirochete Borrelia burgdorferi.";
RL Mol. Microbiol. 35:490-516(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A.,
RA Artiach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
DR EMBL; AE001578; AAF07566.1; -; Genomic_DNA.
DR TIGR; BBM11; -.
DR InterPro; IPR008505; DUF787.
DR Pfam; PF05619; DUF787; 1.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 370 AA; 42158 MW; 62FEE72EB87CF9C3 CRC64;
Query Match 99.4%; Score 1878; DB 2; Length 370;
Best Local Similarity 99.2%; Pred. No. 2.1e-110;
Matches 367; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 MPQDTISVSLDSRIQASRPNNYPLVYTKAKIKVNDAAANYKILNLTNNYKQIETL 60
DB 1 MPQDTISVSLDSRIQASRPNNYPLVYTKAKIKVNDAAANYKILNLTNNYKQIETL 60
QY 61 EKDNNGQDQFGKEKTLTKTAMSNFNSSESLKSADLFYKDKPEELKKYLKVHRHTFV 120
```

```
Db 61 EKONGGQDQFGKTKLLKTAMSNFNSSESLKASADLFIYKDKPEELKKYLVKVRHTFV 120
Qy 121 VLINTEGNSDDGLKIYKDDYKFTPSIFFVSTKEQEIKELFKDKGNTKERNIAVYS 180
Db 121 VLINTEGNSDDGLKIYKDDYKFTPSIFFVSTKEQEIKELFKDKGNTKERNIAVYS 180
Qy 181 NKNKDLHLKFTISQYLHQASIFHAVNPYGMPLAATPLVDDTVIGKLTAKINFYSLNETG 240
Db 181 NKNKDLHLKFTISQYLHQASIFHAVNPYGMPLAATPLVDDTVIGKLTAKINFYSLNETG 240
Qy 241 LDGVPAPKEGVDLAGGAIDQFTYHYIKNEAIIELIRIWNKNRQNSKLSALQLSGARDN 300
Db 241 LDGVPAPKEGVDLAGGAIDQFTYHYIKNEAIIELIRIWNKNRQNSKLSALQLSGARDN 300
Qy 301 AYTSAIECLLKRFDVDRGLIIQYKNLSLTSPQPKLELSVNTYFNFSINAVSLVITTD 360
Db 301 AYTSAIECLLKRFDVDRGLIIQYKNLSLTSPQPKLELSVNTYFNFSINAVSLVITTD 360
Qy 361 IVDYQNSLSA 370
Db 361 IVDYQNSLSA 370

RESULT 4
ID Q9R2W7 BORBU PRELIMINARY; PRT; 370 AA.
AC Q9R2W7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein BBL11 (Hypothetical protein BBL11) (Hypothetical protein BBL11).
GN OrderedLocNames=BBL11, BBL11, BBS11;
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid cp32-8, Plasmid cp32-1, and Plasmid cp32-3.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=139;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31; PLASMID=cp32-8, cp32-1, and cp32-3;
RX MEDLINE=20138354; PubMed=10672174;
RA Casjens S., Palmer N., van Vugt R., Huang W.M., Stevenson B., Rosa P., Lathigra R., Sutton G.G., Peterson J.D., Dodson R.J., Haft D.H., Hickey E.K., Gwinn M.L., White O., Fraser C.M.;
RT "A bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete Borrelia burgdorferi."
RL Mol. Microbiol. 35:490-516(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31; PLASMID=cp32-8, cp32-1, and cp32-3;
RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A., Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K., Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D., Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J., Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A., Artiach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K., Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi."
RL Nature 390:580-586(1997).
DR EMBL; AF001575; AAF07434.1; -; Genomic DNA.
DR EMBL; AF001576; AAF07479.1; -; Genomic DNA.
DR EMBL; AF001580; AAF07651.1; -; Genomic DNA.
DR TIGR; BBL11; -.
DR TIGR; BBL11; -.
DR TIGR; BBS11; -.
DR TIGR; BBS11; -.
DR InterPro; IPR008505; DUF787.
DR Pfam; PF05619; DUF787; 1.
```

```
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 370 AA; 42102 MW; D071E1723FA9767C CRC64;

Query Match 97.2%; Score 1837; DB 2; Length 370;
Best Local Similarity 97.0%; Pred. No. 8e-108;
Matches 359; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MPQDTISVSLDLSRIQASRPYNNPLLVYKTAKIKVNDAAANYKILNLTNNYKQIETL 60
Db 1 MPQDTISVSLDLSRIQASRPYNNPLLVYKTAKIKVNDAAANYKILNLTNNYKQIETL 60
Qy 61 EKONGGQDQFGKTKLLKTAMSNFNSSESLKASADLFIYKDKPEELKKYLVKVRHTFV 120
Db 61 EKONGGQDQFGKTKLLKTAMSNFNSSESLKASADLFIYKDKPEELKKYLVKVRHTFV 120
Qy 121 VLINTEGNSDDGLKIYKDDYKFTPSIFFVSTKEQEIKELFKDKGNTKERNIAVYS 180
Db 121 VLINTEGNSDDGLKIYKDDYKFTPSIFFVSTKEQEIKELFKDKGNTKERNIAVYS 180
Qy 181 NKNKDLHLKFTISQYLHQASIFHAVNPYGMPLAATPLVDDTVIGKLTAKINFYSLNETG 240
Db 181 NKNKDLHLKFTISQYLHQASIFHAVNPYGMPLAATPLVDDTVIGKLTAKINFYSLNETG 240
Qy 241 LDGVPAPKEGVDLAGGAIDQFTYHYIKNEAIIELIRIWNKNRQNSKLSALQLSGARDN 300
Db 241 LDGVPAPKEGVDLAGGAIDQFTYHYIKNEAIIELIRIWNKNRQNSKLSALQLSGARDN 300
Qy 301 AYTSAIECLLKRFDVDRGLIIQYKNLSLTSPQPKLELSVNTYFNFSINAVSLVITTD 360
Db 301 AYTSAIECLLKRFDVDRGLIIQYKNLSLTSPQPKLELSVNTYFNFSINAVSLVITTD 360
Qy 361 IVDYQNSLSA 370
Db 361 IVDYQNSLSA 370

RESULT 5
ID Q9S003 BORBU PRELIMINARY; PRT; 370 AA.
AC Q9S003;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein BBSQ18.
GN OrderedLocNames=BBSQ18;
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=139;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=20138354; PubMed=10672174;
RA Casjens S., Palmer N., van Vugt R., Huang W.M., Stevenson B., Rosa P., Lathigra R., Sutton G.G., Peterson J.D., Dodson R.J., Haft D.H., Hickey E.K., Gwinn M.L., White O., Fraser C.M.;
RT "A bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete Borrelia burgdorferi."
RL Mol. Microbiol. 35:490-516(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A., Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K., Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D., Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J., Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A., Artiach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K., Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
```

"Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.";

RL Nature 390:580-586 (1997).

DR EMBL: AE001584; AAF07741.1; -; Genomic_DNA.

DR TIGR: BBQ18; -.

DR InterPro: IPR008505; DUF787.

DR Pfam: PF05619; DUF787; 1.

KW Complete proteome; Hypothetical protein; Plasmid.

SQ SEQUENCE 370 AA; 42150 MW; DABE039E8F8EBD8 CRC64;

Query Match 96.8%; Score 1829; DB 2; Length 370;

Best Local Similarity 96.8%; Pred. No. 2.6e-107;

Matches 358; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 MPQDTISVSLDSRIQASRPNTYNNPLLYVYKTAKIKVNDAAANYKILNLTVNNYKQIETL 60

Db 1 MPQDTISVSLDSRIQASRPNTYNNPLLYVYKTAKIKVNDAAANYKILNLTVNNYKQIETL 60

QY 61 EKXNGNGDQFGKEKTLTKTAMSNFFNSSESLKSADLFYKDKPEELKCYLKVHRHTFV 120

Db 61 EKXNGNGDQFGKEKTLTKTAMSNFFNSSESLKSADLFYKDKPEELKCYLKVHRHTFV 120

QY 121 VLINTEGNSDGLKIYKDDYDKPTSPISFFVSTKEQEIKELFKDKGNTKERNIAVYS 180

Db 121 VLINTEGNSDGLKIYKDDYDKPTSPISFFVSTKEQEIKELFKDKGNTKERNIAVYS 180

QY 181 NNKDNHLKFIQYHLQASIFHAVNPYGMPLAATPLVDDTVIGKLTAKINFYSLLNETG 240

Db 181 NNKDNHLKFIQYHLQASIFHAVNPYGMPLAATPLVDDTVIGKLTAKINFYSLLNETG 240

QY 241 LDGVPAPFEGVDLAGGADIEQFTYHYIKNEAIIELIRIWNKNRQNSKLSALQSGARDN 300

Db 241 LDGVPAPFEGVDLAGGADIEQFTYHYIKNEAIIELIRIWNKNRQNSKLSALQSGARDN 300

QY 301 AYTSAEICLLKRFVDRGLIIQYKNLSLTSPTPQLKLSVNTYFNFSINAVSLVITQD 360

Db 301 AYTSAEICLLKRFVDRGLIIQYKNLSLTSPTPQLKLSVNTYFNFSINAVSLVITQD 360

QY 361 IVDYQNSLSA 370

Db 361 IVDYQNSLSA 370

RESULT 6

Q9S0F9 BORBU

ID Q9S0F9 BORBU PRELIMINARY; PRT; 370 AA.

AC Q9S0F9;

DT 01-MAY-2000 (TReMBLrel. 13, Created)

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)

DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)

DE Hypothetical protein BBR11.

GN OrderedLocusNames=BBR11;

OS Borrelia burgdorferi (Lyme disease spirochete).

OC Bacteri; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;

OC Bacteri; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;

OX NCBI_TaxID=139;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=ATCC 35210 / B31;

RX MEDLINE=20138354; PubMed=10672174;

RA Casjens S., Palmer N., van Vugt R., Dodson R.J., Haft D.H.,

RA Lathigra R., Sutton G.G., Peterson J.D., White O., Fraser C.M.;

RA Hickey E.K., Gwinn M.L., White O., Fraser C.M.;

RT "A bacterial genome in flux: the twelve linear and nine circular

RT extrachromosomal DNAs in an infectious isolate of the Lyme disease

RT spirochete Borrelia burgdorferi.";

RL Mol. Microbiol. 35:490-516 (2000).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=ATCC 35210 / B31;

RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;

RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,

RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,

RA Gwinn M.L., Dougherty B.A., Tomb J.-P., Fleischmann R.D.,

RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,

RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,

RA Gocayne J.D., Weidman J.F., Uterback T.R., Wathey L.D., McDonald L.A.,

RA Artiach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,

RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;

RT "Genomic sequence of a Lyme disease spirochaete, Borrelia

RT burgdorferi.";

RL Nature 390:580-586 (1997).

DR EMBL: AE001577; AAF07520.1; -; Genomic_DNA.

DR TIGR: BBR11; -.

DR InterPro: IPR008505; DUF787.

DR Pfam: PF05619; DUF787; 1.

KW Complete proteome; Hypothetical protein; Plasmid.

SQ SEQUENCE 370 AA; 42130 MW; BE37D9DB14A1B0BD CRC64;

Query Match 96.8%; Score 1828; DB 2; Length 370;

Best Local Similarity 96.2%; Pred. No. 3e-107;

Matches 356; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 1 MPQDTISVSLDSRIQASRPNTYNNPLLYVYKTAKIKVNDAAANYKILNLTVNNYKQIETL 60

Db 1 MPQDTISVSLDSRIQASRPNTYNNPLLYVYKTAKIKVNDAAANYKILNLTVNNYKQIETL 60

QY 61 EKXNGNGDQFGKEKTLTKTAMSNFFNSSESLKSADLFYKDKPEELKCYLKVHRHTFV 120

Db 61 EKXNGNGDQFGKEKTLTKTAMSNFFNSSESLKSADLFYKDKPEELKCYLKVHRHTFV 120

QY 121 VLINTEGNSDGLKIYKDDYDKPTSPISFFVSTKEQEIKELFKDKGNTKERNIAVYS 180

Db 121 VLINTEGNSDGLKIYKDDYDKPTSPISFFVSTKEQEIKELFKDKGNTKERNIAVYS 180

QY 181 NNKDNHLKFIQYHLQASIFHAVNPYGMPLAATPLVDDTVIGKLTAKINFYSLLNETG 240

Db 181 NNKDNHLKFIQYHLQASIFHAVNPYGMPLAATPLVDDTVIGKLTAKINFYSLLNETG 240

QY 241 LDGVPAPFEGVDLAGGADIEQFTYHYIKNEAIIELIRIWNKNRQNSKLSALQSGARDN 300

Db 241 LDGVPAPFEGVDLAGGADIEQFTYHYIKNEAIIELIRIWNKNRQNSKLSALQSGARDN 300

QY 301 AYTSAEICLLKRFVDRGLIIQYKNLSLTSPTPQLKLSVNTYFNFSINAVSLVITQD 360

Db 301 AYTSAEICLLKRFVDRGLIIQYKNLSLTSPTPQLKLSVNTYFNFSINAVSLVITQD 360

QY 361 IVDYQNSLSA 370

Db 361 IVDYQNSLSA 370

RESULT 7

Q9S0S1 BORBU

ID Q9S0S1 BORBU PRELIMINARY; PRT; 362 AA.

AC Q9S0S1;

DT 01-MAY-2000 (TReMBLrel. 13, Created)

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)

DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)

DE Hypothetical protein BBN11.

GN OrderedLocusNames=BBN11;

OS Borrelia burgdorferi (Lyme disease spirochete).

OC Bacteri; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;

OC Bacteri; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;

OX NCBI_TaxID=139;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=ATCC 35210 / B31;

RX MEDLINE=20138354; PubMed=10672174;

RA Casjens S., Palmer N., van Vugt R., Dodson R.J., Haft D.H.,

RA Lathigra R., Sutton G.G., Peterson J.D., White O., Fraser C.M.;

RA Hickey E.K., Gwinn M.L., White O., Fraser C.M.;

RT "A bacterial genome in flux: the twelve linear and nine circular

RT extrachromosomal DNAs in an infectious isolate of the Lyme disease


```

RT spirochete Borrelia burgdorferi.;
RL Mol. Microbiol. 35:490-516(2000).
RN [2]
RC NUCLEOTIDE SEQUENCE.
RX STRAIN=ATCC 35210 / B31;
RM MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Frazer C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M.J., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Weidman J.F., Utterback T.R., Wathey L., McDonald L.A.,
RA Artiach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
DR EMBL; AB001581; AAF07685.1; -; Genomic_DNA.
DR TIGR; BBN11; -;
DR InterPro; IPR008505; DUF787.
DR Pfam; PF05619; DUF787; 1.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 362 AA; 41367 MW; CD76E4F607088587 CRC64;

Query Match 92.9%; Score 1755; DB 2; Length 362;
Best Local Similarity 94.2%; Pred. No. 1.2e-102;
Matches 341; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

Qy 1 MPQDTISVSLDSRIQASRPYNYPLLYVTAKIKVNDAAANYKILNLTNNYEQIETL 60
Db 1 MPQDTISVSLDSRIQASRPYNYPLLYVTAKIKVNDAAANYKILNLTNNYEQIETL 60
Qy 61 EKONGGODQFGKEKTLTKTAMSNFNSSESLKADLFYKDKPEELKYLKVRHTFV 120
Db 61 EKONGGODQFGKEKTLTKTAMSNFNSSESLKADLFYKDKPEELKYLKVRHTFV 120
Qy 61 EKENGNEEDQFGKEKTLTKTAMSNFNSSESLKADLFYKDKPEELKYLKVRHTFV 120
Db 61 EKENGNEEDQFGKEKTLTKTAMSNFNSSESLKADLFYKDKPEELKYLKVRHTFV 120
Qy 121 VLINTEGNSDGLKIYKDDYDKFTPSIFFVFSTKEQEIKELFKD-KGNTKERNIA 177
Db 121 VLINTEGNSDGLKIYKDDYDKFTPSIFFVFSTKEQEIKELFKD-KGNTKERNIA 177
Qy 121 VLINTEGNSDGLKIYKDDYDKFTPSIFFVFSTKEQEIKELFKD-KGNTKERNIA 177
Db 121 VLINTEGNSDGLKIYKDDYDKFTPSIFFVFSTKEQEIKELFKD-KGNTKERNIA 177
Qy 178 VYSNNKNLHLKFTISQYLHQASIFHAVNPYGMPLAATPLVDVTGKLTAKINFYSLN 237
Db 178 VYSNNKNLHLKFTISQYLHQASIFHAVNPYGMPLAATPLVDVTGKLTAKINFYSLN 237
Qy 238 ETGLDGPAPKEGVDLAGAIDEQFTHYIKNEAIIELIRIWNKNRNSKLSALQSLGA 297
Db 238 ETGLDGPAPKEGVDLAGAIDEQFTHYIKNEAIIELIRIWNKNRNSKLSALQSLGA 297
Qy 241 ETGLDGPAPKEGVDLAGAIDEQFTHYIKNEAIIELIRIWNKNRNSKLSALQSLGA 300
Db 241 ETGLDGPAPKEGVDLAGAIDEQFTHYIKNEAIIELIRIWNKNRNSKLSALQSLGA 300
Qy 298 RDNATSAIECLLKRFDVDRGLIIQYKNLSLTSPTPQLKLEISVNTYFNFSINAVSLVIT 357
Db 298 RDNATSAIECLLKRFDVDRGLIIQYKNLSLTSPTPQLKLEISVNTYFNFSINAVSLVIT 357
Qy 301 RDNATSAIECLLKRFDVDRGLIIQYKNLSLTSPTPQLKLEISVNTYFNFSINAVSLVIT 360
Db 301 RDNATSAIECLLKRFDVDRGLIIQYKNLSLTSPTPQLKLEISVNTYFNFSINAVSLVIT 360
Qy 358 TQDIVDYQNSL 368
Db 358 TQDIVDYQNSL 368
Qy 361 SODIVDYQNSL 371
Db 361 SODIVDYQNSL 371

RESULT 9
O50939 BORBU PRELIMINARY; PRT; 356 AA.
ID O50939 BORBU PRELIMINARY; PRT; 356 AA.
AC O50939;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein BBA46.
GN OrderedLocusNames=BEA46;
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp54.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=139;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Frazer C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M.J., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Weidman J.F., Utterback T.R., Wathey L., McDonald L.A.,
RA Artiach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
DR EMBL; AB001581; AAF07685.1; -; Genomic_DNA.
DR TIGR; BBN11; -;
DR InterPro; IPR008505; DUF787.
DR Pfam; PF05619; DUF787; 1.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 362 AA; 41367 MW; CD76E4F607088587 CRC64;

Query Match 92.9%; Score 1755; DB 2; Length 362;
Best Local Similarity 94.2%; Pred. No. 1.2e-102;
Matches 341; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

Qy 1 MPQDTISVSLDSRIQASRPYNYPLLYVTAKIKVNDAAANYKILNLTNNYEQIETL 60
Db 1 MPQDTISVSLDSRIQASRPYNYPLLYVTAKIKVNDAAANYKILNLTNNYEQIETL 60
Qy 61 EKONGGODQFGKEKTLTKTAMSNFNSSESLKADLFYKDKPEELKYLKVRHTFV 120
Db 61 EKONGGODQFGKEKTLTKTAMSNFNSSESLKADLFYKDKPEELKYLKVRHTFV 120
Qy 61 EKENGNEEDQFGKEKTLTKTAMSNFNSSESLKADLFYKDKPEELKYLKVRHTFV 120
Db 61 EKENGNEEDQFGKEKTLTKTAMSNFNSSESLKADLFYKDKPEELKYLKVRHTFV 120
Qy 121 VLINTEGNSDGLKIYKDDYDKFTPSIFFVFSTKEQEIKELFKD-KGNTKERNIA 177
Db 121 VLINTEGNSDGLKIYKDDYDKFTPSIFFVFSTKEQEIKELFKD-KGNTKERNIA 177
Qy 121 VLINTEGNSDGLKIYKDDYDKFTPSIFFVFSTKEQEIKELFKD-KGNTKERNIA 177
Db 121 VLINTEGNSDGLKIYKDDYDKFTPSIFFVFSTKEQEIKELFKD-KGNTKERNIA 177
Qy 178 VYSNNKNLHLKFTISQYLHQASIFHAVNPYGMPLAATPLVDVTGKLTAKINFYSLN 237
Db 178 VYSNNKNLHLKFTISQYLHQASIFHAVNPYGMPLAATPLVDVTGKLTAKINFYSLN 237
Qy 238 ETGLDGPAPKEGVDLAGAIDEQFTHYIKNEAIIELIRIWNKNRNSKLSALQSLGA 297
Db 238 ETGLDGPAPKEGVDLAGAIDEQFTHYIKNEAIIELIRIWNKNRNSKLSALQSLGA 297
Qy 241 ETGLDGPAPKEGVDLAGAIDEQFTHYIKNEAIIELIRIWNKNRNSKLSALQSLGA 300
Db 241 ETGLDGPAPKEGVDLAGAIDEQFTHYIKNEAIIELIRIWNKNRNSKLSALQSLGA 300
Qy 298 RDNATSAIECLLKRFDVDRGLIIQYKNLSLTSPTPQLKLEISVNTYFNFSINAVSLVIT 357
Db 298 RDNATSAIECLLKRFDVDRGLIIQYKNLSLTSPTPQLKLEISVNTYFNFSINAVSLVIT 357
Qy 301 RDNATSAIECLLKRFDVDRGLIIQYKNLSLTSPTPQLKLEISVNTYFNFSINAVSLVIT 360
Db 301 RDNATSAIECLLKRFDVDRGLIIQYKNLSLTSPTPQLKLEISVNTYFNFSINAVSLVIT 360
Qy 358 TQDIVDYQNSL 368
Db 358 TQDIVDYQNSL 368
Qy 361 SODIVDYQNSL 371
Db 361 SODIVDYQNSL 371

RESULT 8
Q5XZ68 BORG
ID Q5XZ68 BORG PRELIMINARY; PRT; 373 AA.
AC Q5XZ68;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=BGP064;
OS Borrelia garinii PBI.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=290434;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBI;

```

RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi";
RL Nature 390:580-586(1997).
DR EMBL; AE000790; AAC62278.1; -; Genomic_DNA.
DR FIR; F70212; F70212.
DR TIGR; BBA46; -.
DR InterPro; IPR008505; DUF787.
DR Pfam; PF05619; DUF787; 1.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 356 AA; 40359 MW; AL1343C1F56ED474 CRC64;

Query Match 42.5%; Score 803.5; DB 2; Length 356;
Best Local Similarity 46.3%; Pred. No. 1.2e-42;
Matches 171; Conservative 58; Mismatches 117; Indels 23; Gaps 6;

QY 1 MPQDTISVSLDSRIQASRPNTYNNPLLYVTAKIKVNDAAANYKILNLTNNYKQIETL 60
DB 1 MPKDTISVSVQNLAFNVNTYNNPLLYYKS-----NVLASGHLASVTNLODLEKL 53
QY 61 EKDNNGGQDQFGK----EKTLLKTAMSNFFNSSESLKSADLFYKDKPEELKYLKVH 115
DB 54 ENQKAQESDSKKTAEBQLSKMSDF--GDGLKSVDFYVI-NQIKEIKDFLKS 110
QY 116 RHTFVVLINTEGNSDDGLKIYKDDYDKFKTPSIFVFSTKEQEIKELFKDKGNTKERN 175
DB 111 LHPFVVVFNQAE-----VISSDFEIRKICNFIVISTKDANLPDLFGKDKSELKNI 163
QY 176 IAVYNNKDNHLKIFISVYLHOASIFHAWNPGMPLAATPLVDDTVIGKLTAKINFYSL 235
DB 164 IAVYA-GKQNLHLKFTAAVYLHOASIFHAVNPYGMILNAVVPVYDDSLDLSLRKTINFYSL 222
QY 236 LNETGLDGVPAFKEGVDLAGGAI DEQTVHYIKNEAIIELIRIWNKNRQNSKLSALQLS 295
DB 223 LNETGNDGLAFKEGVSLSGDPIDETFTLYIKNEAKELIRIWNKNRANSKLSGLND 282
QY 296 GARDNAYTSAIECLLKRFDVDRGLLIQYKNLSLTLSPTQKLELSVNTIYFNSINAVSIV 355
DB 283 GNLNPEYTAGLCFFHELKQGLIVSYKEIKLINSSEGLSLSLEVALKYNDSPNSVNI 342
QY 356 ITQDIDYV 364
DB 343 ITTOEINEY 351

RESULT 10
QASIV3_BORGA PRELIMINARY; PRT; 106 AA.
AC Q5XYV3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=BGPI78;
OS Borrelia garinii pBi.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=290434;
RN [1]
RP NUCLEOTIDE SEQUENCE.

Query Match 20.2%; Score 382.5; DB 2; Length 156;
Best Local Similarity 49.1%; Pred. No. 1.5e-16;
Matches 82; Conservative 24; Mismatches 50; Indels 11; Gaps 4;

QY 82 MSNFFNSSESLKSADLFYKDKPEELKYLKVHRTFVVLINTEGNSDDGLKIYKDDY 141
DB 1 MGDFF--GDGLKSVDFYVI-NQIKEIKDFLKSNLHPFVFLNESGD-----ALSSDF 50
QY 142 DKFKTPSIFVFSTKEQEIKELFKDKGNTKERNIATYVNNKDNHLKFIISOYLHOASIF 201
DB 51 EAIRKALNFIVISTKENGFLNFKGDKGELKNIATYVSGN-ENLHLKFAAMYLHOASIF 109
QY 202 HAVNPYGMPLAATPLVDDTVIGKLTAKINFYSLINTEGNDGLGVPAFK 248
DB 110 HAVNPYGMPLNSTFYDDSLDLSLRKANINFYSLINTEGNDGLAFK 156

RESULT 12
QASIV2_BORGA PRELIMINARY; PRT; 76 AA.
AC Q6ASII;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BGA43;
OS Borrelia garinii.
OX Plasmid lp54.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.

KW Hypothetical protein.
SQ SEQUENCE 106 AA; 11884 MW; DOA7E975F9574FB2 CRC64;

Query Match 22.3%; Score 421; DB 2; Length 106;
Best Local Similarity 83.0%; Pred. No. 3.6e-19;
Matches 83; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

QY 1 MPQDTISVSLDSRIQASRPNTYNNPLLYVTAKIKVNDAAANYKILNLTNNYKQIETL 60
DB 1 MPKDTISVSVQNLAFNVNTYNNPLLYYKS-----NVLASGHLASVTNLODLEKL 53
QY 61 EKDNNGGQDQFGKERTLLKTAMSNFFNSSESLKSADLF 100
DB 61 EKDNNGSNDQLGKEGTLLKTAMSNFFNSTEESLSKSAVLFI 100

RESULT 11
QASII_BORGA PRELIMINARY; PRT; 156 AA.
AC Q6ASII;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BGA44;
OS Borrelia garinii.
OX Plasmid lp54.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=29519;
RN [1]
RP NUCLEOTIDE SEQUENCE.

Query Match 20.2%; Score 382.5; DB 2; Length 156;
Best Local Similarity 49.1%; Pred. No. 1.5e-16;
Matches 82; Conservative 24; Mismatches 50; Indels 11; Gaps 4;

QY 82 MSNFFNSSESLKSADLFYKDKPEELKYLKVHRTFVVLINTEGNSDDGLKIYKDDY 141
DB 1 MGDFF--GDGLKSVDFYVI-NQIKEIKDFLKSNLHPFVFLNESGD-----ALSSDF 50
QY 142 DKFKTPSIFVFSTKEQEIKELFKDKGNTKERNIATYVNNKDNHLKFIISOYLHOASIF 201
DB 51 EAIRKALNFIVISTKENGFLNFKGDKGELKNIATYVSGN-ENLHLKFAAMYLHOASIF 109
QY 202 HAVNPYGMPLAATPLVDDTVIGKLTAKINFYSLINTEGNDGLGVPAFK 248
DB 110 HAVNPYGMPLNSTFYDDSLDLSLRKANINFYSLINTEGNDGLAFK 156

```

OX NCBI_TaxID=29519;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PBI;
RA Gloeckner G., Lehmann R., Romualdi A., Pradella S.,
RA Schulte-Spechtel U., Wilske B., Suehnell J., Platzer M.;
RT "Comparative analysis of the Borrelia garinii genome.";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP0000015; AAT93802.1; -; Genomic_DNA.
DR InterPro; IPR008505; DUF787.
DR Pfam; PF05619; DUF787; 1.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 76 AA; 8637 MW; 63268818A20B5290 CRC64;

Query Match 7.4%; Score 140.5; DB 2; Length 76;
Best Local Similarity 41.9%; Pred. No. 0.12;
Matches 31; Conservative 15; Mismatches 21; Indels 7; Gaps 1;

QY 1 MPQDTISVLSLDSRIQASRPYNNPLVYTKAKVKNDAANYKILNLTVNNYEKQIETL 60
Db 1 MPQDTISVLSLDSRIQASRPYNNPLVYTKAKVKNDAANYKILNLTVNNYEKQIETL 60
QY 61 EKONGGQDQFGKE 74
Db 54 EIRKGETDSSNKK 67

RESULT 13
QYVZ02 CRYPV PRELIMINARY; PRT; 2670 AA.
AC QYVZ02;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Hypothetical protein.
GN ORFNames=IWS.772;
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=5807;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Bankier A.T., Spriggs H.F., Fartmann B., Konfortov B.A., Madera M.,
RA Vogel C., Teichmann S.A., Ivens A., Dear P.H.;
RT "Integrated mapping, chromosomal sequencing and sequence analysis of
RT Cryptosporidium parvum.";
RL Genome Res. 0:0-0(2003).
DR EMBL; BX538353; CAD98328.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 2670 AA; 317723 MW; 670B6DF63D01A83B CRC64;

Query Match 7.4%; Score 140; DB 2; Length 2670;
Best Local Similarity 21.2%; Pred. No. 9.7;
Matches 93; Conservative 77; Mismatches 130; Indels 138; Gaps 22;

QY 3 QDTISVLSLDSRIQASRPYNNPLVYTKAKVKNDAANYKILNLTVNNYEKQIETLEK 62
Db 1006 QESIKI-----IKSKDPNIKN--LIQKDDIKISKERN--IVNL-VNRYLSIIE---- 1050

QY 63 DNGNGQDQKGKTKLTAMSNF-----FNSSSELSKSADLFYKPKBELKKYLVKH 115
Db 1051 DERN-----QILLILGILSYRISRMNLNENELLKFNLF--KDLIEK-KKVITIN 1100

QY 116 RHTFVVLIN-----TEGNSDDGLK-IYKODYDKFKTPSPFFVPS 154
Db 1101 SKLFGYLNDWIHFISFIINTARNQIDMNDNHHFLKEWFNDNLYRKYTP--YKIFN 1158

QY 155 TKEQETK-----ELPKDKGNTKERN-----IAY 178
Db 1159 SEKILEKKNVSLILFKLIKNIKMKNLVNNDKNVLEIVNTRENDKSDFFYNVLV 1218

QY 179 YGNKNLHLKFIQSLVHQASIFHAVNYPGMPAAATPLVDVDTVIGKLTAKINFYSLNE 238
Db 1031 QESIKI-----IKSKDPNIKN--LIQKDDIKISKERN--IVNL-VNRYLSIIE---- 1075

```

```

QY 63 DNGNGQDFGKKTLLKTSMSNF-----FNSSESLSKADLFYKDKPEELKYLKVH 115
Db 1076 DEEN-----QILLGLLSYSIRISNMLNELLKFNLF--KLLIEK-KKVIIN 1125
QY 116 RHTFVLIN-----TEGNSDDGLK-IYKDDYDKFTPSIFFVFS 154
Db 1126 SKLFGYLNDWTHFEISFIINARMQNDMNDNHFILKEWFDNLYRYKTP--YKIFN 1183
QY 155 TKEQIK-----ELFKDKGNTKERN-----TAV 178
Db 1184 SCEKILEKNENVSLLIFKLIRKISMKYLNVLNNDKNYLEIVKNTRENDKGRDFYNVLV 1243
QY 179 YSNKNLHLKFILOYHQAISFHAVNPGMPLAATPLVDVDTVIGKLTAKINFYSLLNE 238
Db 1244 YNQS-----VIRRIQGLDSTINRF-----NNLSENEICQMKSERIK----- 1283
QY 239 TGLDGVPAFKEGVDLAGGAIDQFTYHYIKNBAILIELIRIMNKNRQNSK-----LSALQL 294
Db 1284 -----EYQRLRGLFGDLMPLPKDYNNETFDIVRNGUYNMRTSSKFLAVYPPSILI 1334
QY 295 SCARDNATSAIECLIKRPVDRGLIIOY-KNLSLTSPQKLELSVNIITN-----FSIN 350
Db 1335 ESLRKEPENLSCGNPLRPIEBISMYFSKDLALL-----VKINMTYKCBIFIID 1385
QY 351 AVSLVITTDIVDYQNSL 368
Db 1386 A-ELLITLSSEIELKSSI 1402

RESULT 15
Q7RB65 PLAYO
ID Q7RB65 PLAYO PRELIMINARY; PRT; 1057 AA.
AC Q7RB65;
DT 01-WAR-2004 (TremBLrel. 26, Created)
DT 01-WAR-2004 (TremBLrel. 26, Last sequence update)
DT 01-WAR-2004 (TremBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=PY06283;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=17XNL;
RX MEDLINE=22255706; PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Perteaux M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallow S.J., van Aken S.E., Riedmuller S.B., Feldblum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabai A., Cummings L.M.,
RA Florens L., Yates J.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii."
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AB01002103; EAA18474.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 1057 AA; 125585 MW; 9C4CB84F0A11661 CRC64;

Query Match 7.3%; Score 137; DB 2; Length 1057;
Best Local Similarity 23.5%; Pred. No. 4.8;
Matches 97; Conservative 46; Mismatches 151; Indels 118; Gaps 21;

QY 21 NYTNPLLVYKAKIKVKNDAANY-----KILMLTVNNYKQIETLEKON----- 64
Db 171 NYVKD-IIFKFLPKENNITQLKPNKLPFLFSLNLSHDGNDKMYQIFKKKNPKFNSQLL 229

```

```

QY 65 ---GNGDQFGK---EKTLLKT---AMNPFNSSESLSKADLFYKDKPEELKYLKVHRH 117
Db 230 SSVKNGIDGFLKYPKTKTVIDTHGIYGNPLTYEYTKKABIFRPNKEEDHLNPF-DIFLS 288
QY 118 TFVVLINTE-----GNSDDGLKIYKD-----DYDKFKTPSIFPFVSTKEQ 158
Db 289 FFIFISINTEIRKHLKHQVSDVDLCKLKHDPHFNFVDLIPHIDYEHIN--DVFFRFM--- 342
QY 159 EIKELFKDKGNTKERNIA-----VYSNNKDNLHLKFILOYHQAISFHAVNPGM 209
Db 343 -----NTTKTHNILPLMLFLRFFHPKND-----LFTYSNVINAVEKY-- 381
QY 210 PLAATPLVDVDT--VIGKLTAKINFYSLLNETGLDGVPAFKEGVD--LAGGAIDQFTVH 265
Db 382 -FTSEIFDSDTKKMIGKLA-----YSLMARLSFTITVSYLEPKIDDDVIDGIDIDLMNLFH 434
QY 266 YIKNEAIIEL-----IRWNKNRQNSKLSALQ-----LSGARDNATSAIEC 308
Db 435 NVINKMVVTLCKEDKKLIEFPFVPISMVKNKMKFKLDALKIYNBFTITSLPDQDKESAINF 494
QY 309 LLKRP--VDRGLI---IQYKNLSLTSP-----TPOLKLELSVNIITNPSIN 350
Db 495 LQTDPPKVVVTTLISDALDFVNRHLEKGPDTYHSDEYVKLETFLFKINYQISSN 546

Search completed: January 24, 2006, 19:57:00
Job time : 81.505 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:30:59 ; Search time 22.8097 Seconds
(without alignments)
1341.095 Million cell updates/sec

Title: US-10-688-058-20

Perfect score: 1889

Sequence: 1 MPQDTISVLLDSRIQASRP.....AVSLVITQDIVDYQNSLSA 370

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5 COMB.pep:*
- 2: /cgn2_6/ptodata/1/iaa/6 COMB.pep:*
- 3: /cgn2_6/ptodata/1/iaa/H-COMB.pep:*
- 4: /cgn2_6/ptodata/1/iaa/PCRTUS COMB.pep:*
- 5: /cgn2_6/ptodata/1/iaa/RE COMB.pep:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	132.5	7.0	1155	2	US-09-710-279-1780
2	130.5	6.9	1742	2	US-09-386-962C-4
3	130.5	6.9	1742	2	US-09-386-959-4
4	119	6.3	612	2	US-09-538-092-775
5	116	6.1	524	2	US-09-543-681A-4787
6	112.5	6.0	340	2	US-09-949-016-10217
7	112.5	6.0	382	1	US-08-477-451-28
8	111	5.9	1087	2	US-09-830-230A-322
9	111	5.9	1119	2	US-09-830-230A-321
10	111	5.9	1241	2	US-09-248-796A-18397
11	110	5.8	1078	2	US-09-248-796A-20284
12	109	5.8	1164	2	US-08-923-992A-10
13	107	5.7	429	2	US-09-134-000C-5630
14	107	5.7	544	2	US-09-710-279-38
15	105	5.6	1060	2	US-09-248-796A-15476
16	105	5.6	1098	2	US-08-923-992A-8
17	104.5	5.5	400	2	US-09-248-796A-15785
18	104.5	5.5	458	2	US-09-830-230A-30
19	104.5	5.5	481	2	US-09-830-230A-29
20	104.5	5.5	561	2	US-09-248-796A-24349
21	104	5.5	1164	2	US-08-923-992A-2
22	103.5	5.5	491	2	US-09-029-267-2
23	103.5	5.5	677	2	US-09-543-681A-5460
24	103.5	5.5	691	2	US-09-538-092-421
25	103.5	5.5	3878	2	US-09-914-259-11
26	103	5.5	1781	1	US-08-477-451-11
27	102	5.4	488	2	US-09-248-796A-18820

28	102	5.4	3135	1	US-08-323-170B-2	Sequence 2, Appli
29	102	5.4	3135	2	US-08-954-441-2	Sequence 2, Appli
30	101.5	5.4	439	2	US-09-538-092-1216	Sequence 1216, Ap
31	101.5	5.4	736	2	US-09-583-110-4555	Sequence 4555, Ap
32	101.5	5.4	741	2	US-09-107-433-4401	Sequence 4401, Ap
33	101.5	5.4	773	2	US-09-841-786-6	Sequence 6, Appli
34	101.5	5.4	3241	2	US-09-841-786-1	Sequence 1, Appli
35	101	5.3	820	2	US-09-248-796A-19901	Sequence 19901, A
36	101	5.3	912	1	US-08-951-871-2	Sequence 2, Appli
37	101	5.3	972	2	US-09-248-796A-20234	Sequence 20234, A
38	100.5	5.3	2184	2	US-09-417-485D-6	Sequence 6, Appli
39	100.5	5.3	2710	1	US-08-480-604A-6	Sequence 6, Appli
40	100.5	5.3	2710	2	US-08-405-496A-6	Sequence 6, Appli
41	100.5	5.3	2710	2	US-08-915-136-6	Sequence 6, Appli
42	100.5	5.3	2710	2	US-08-957-310-6	Sequence 6, Appli
43	100.5	5.3	2710	2	US-10-011-366-6	Sequence 6, Appli
44	100.5	5.3	2710	2	US-09-084-517-6	Sequence 6, Appli
45	100	5.3	717	2	US-09-107-532A-6277	Sequence 6277, Ap

ALIGNMENTS

RESULT 1

US-09-710-279-1780

; Sequence 1780, Application US/09710279

; Patent No. 6703492

; GENERAL INFORMATION:

; APPLICANT: KIMMERLY, WILLIAM JOHN

; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

; FILE REFERENCE: PU3480US

; CURRENT APPLICATION NUMBER: US/09710,279

; CURRENT FILING DATE: 2000-11-09

; PRIOR APPLICATION NUMBER: 60/164,258

; PRIOR FILING DATE: 1999-11-09

; NUMBER OF SEQ ID NOS: 4472

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1780

; LENGTH: 1155

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic

; OTHER INFORMATION: amino acid sequence

US-09-710-279-1780

Query Match						7.0%; Score 132.5; DB 2; Length 1155;
Best Local Similarity						22.3%; Pred. No. 0.0048;
Matches						96; Conservative 58; Mismatches 169; Indels 107; Gaps 21;
Qy	18	SRPNYPLLVYKTAIKVKV	DAANYKILMLTVNVEKQIETLEKONGQDQFGKEKT	76		
Db	113	SKPRANEAVTNSTKPKTTEAPTVEESIAETPKTSTTQDSTKKNPSLKNLSST	172			
Qy	77	LLKTAMNFFNSSSESLSADLFYK	-----DKPEELKKYLKVHRHFFVLVINTEGNS	130		
Db	173	TSKES	-----KTDEHSTKQACMTKSNLNDTNDSTPQSEK	220		
Qy	131	D	-----DGLKIVKDDYD	-----KFKTPSIFVFFSTKEQS	162	
Db	221	SAPSKQSDSKSEQKVYKTFNDEPTQDVEHTTKLTPSISTDSSVNDKQDYTRSAVAS	280			
Qy	163	LPKDKGNTKEKRNLAIVSNKKNLHLKFPIS	QYLHQASIFHAV	-----NP-YGM	213	
Db	281	LGVSNETE	-----AITNAVDNLDLKAASREQINEALIAALKKDFSNPDYGVDTPLAL	335		
Qy	214	TPLVDDTVIGKLTAKINFYSLNLTGL	-----DCVPAPKEGVDL	-----AGGAI	258	
Db	336	NTSQSKSPHKSASPRMNLMSLAAEPNSGKNVDKVKITNPTLSLNKNHANNVWPTS	395			
Qy	259	DEQFTY	-----HYTKNEAI	-----IELIRIWNKNRQNSKLSALQSGAR	298	
Db	396	NEQFNLKANYELDDSIKEGDTFTIKYQYIRPGGLELPAI	-----KTQLRSKQGSIVANGVY	452		


```
; LENGTH: 612
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)-(0)
; OTHER INFORMATION: Polypeptide Accession Number YPL070W
US-09-538-092-775

Query Match          6.3%; Score 119; DB 2; Length 612;
Best Local Similarity 21.3%; Pred. No. 0.032;
Matches 86; Conservative 58; Mismatches 138; Indels 122; Gaps 20;

Qy 19 RPNYPLVYKTAKI-----KVNDAANYKIINLTNN--YEKQIETLEKNGNGQDQFG 72
Db 102 QPKVRLVQLSGLFQSFYIKDKASFQHVSSANNNGY-----PSGGSSSFL 152

Qy 73 KEKTLTKAMSNFNSSES-----LKSADLFYKD-----KPELKKYLK 113
Db 153 AAKETLSSGLSGIFGRSSSSGNSLMRPRSSSLFSNESISNSTNATQMLSPKIKQLK 212

Qy 114 VHRFTFVLINTEGNSDGLKIYK---DDYDKFTPSIFFVSTKEQIKELFKDGN 169
Db 213 I-----NELNMKIEKMBELCERDVPK--KILIV----- 239

Qy 170 TEKERNIAVYNNKDNLHKITSQYLHQASIFHAVNPGMPLAAT---PLVDDTV--IGK 224
Db 240 -----GTSVSPNK-----MKTFPK--HQLOTFKVGNLFSNVEFTYNNKLNKILCLSK 288

Qy 225 LRTA-KINFYSLN-ETGLDGPAPKEGVDLAGAIDEQFTYHYI---KNEAIELIRI 278
Db 289 LSTWKNILKFLSLNNGIDPEKPEIKDIL-----YEFTVHSISPCEKIKALLKHEI 343

Qy 279 WKNKRNQSKLSALQSGARDNAYTSAIECLLKRFDVRL-----IIQYKNLSLTS 330
Db 344 MTYSQEMS-----NDDYLSLIYIITVPRDIFLNAEFIRLFYKPKLVETE 391

Qy 331 PTPQLKELSV-----NITVNSINAVSLVITQDIVDYQNSLSA 370
Db 392 SFALTNLKAALFVVEGLTNDFSNELQDKLTVNESKILENSISS 435

RESULT 5
US-09-543-681A-4787
; Sequence 4787, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4787
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4787

Query Match          6.1%; Score 116; DB 2; Length 524;
Best Local Similarity 19.7%; Pred. No. 0.048;
Matches 71; Conservative 51; Mismatches 144; Indels 94; Gaps 12;

Qy 1 MPQDTISVLSLDSRTOASRPYNNPLVYKTAKIKVNDAANYKIINLTNNYKQIE-- 58
Db 245 LENDINLNDINLNDINLNDINLNDINLNDINLNDINLNDINLNDINLNDINLNDIN 304

Qy 59 --TLEKNGNGDQKGKTKLTKTAMSN--PFNSSESLSADL-----FYKDKPEEL 108
Db 305 NINLENDNAN-----LNKIKALIKDNRNTYIFNGKEIKSQODKAMKGFETLTKDIS 360

; LENGTH: 612
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)-(0)
; OTHER INFORMATION: Polypeptide Accession Number YPL070W
US-09-538-092-775

Query Match          6.0%; Score 112.5; DB 2; Length 340;
Best Local Similarity 20.3%; Pred. No. 0.052;
Matches 60; Conservative 45; Mismatches 101; Indels 89; Gaps 13;

Qy 25 PLLVYKTAKIKVNDAANYKIINLTNNYKQIETLEKNGNGDQDQFGKTKLTKTAMSN 84
Db 102 PVFIRKDRVTVO-----EREAEAL-----KQKELEQEA-- 131

Qy 85 PFNSSESLSADLFIYKDKPEELKKYLKVRHRTFVVL--INTGNSDDGLKIYKDYD 142
Db 132 --RMAEERRK---YTLKIVEETKKELENKRSIALDALNTDDEND-----EVEE 178

Qy 143 KFKTPSIFVSTKEQIKELFKDKNTEKERNIA-----VYSNNKDNLHLK 189
Db 179 AWKVRCLKRIKRDREDR-EALEKEKAEIERMNLTEERRAEALRANGKVIITNKAVGKYK 237

Qy 190 FISQYLHQASIF--HAVNPGMPLAATPLVD---DTVIGKLTAKINFYSLNLTGLDV 244
Db 238 FLQKYTHRGAFPMDEDEEVYKRDPSAPTELDHFNKTILPKVNVQV--NF-----GRSGR 289

Qy 245 PAFKEGVDLAGAIDEQFTYHYIKNEAIELIRIWNKRNQSKLSALQLSGARD 299
Db 290 TKYTHLVDDTTSFDS-----AWQBSAQNTKFKKQKAAGVRD 327

RESULT 6
US-09-949-016-10217
; Sequence 10217, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10217
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10217

Query Match          6.0%; Score 112.5; DB 2; Length 340;
Best Local Similarity 20.3%; Pred. No. 0.052;
Matches 60; Conservative 45; Mismatches 101; Indels 89; Gaps 13;

Qy 25 PLLVYKTAKIKVNDAANYKIINLTNNYKQIETLEKNGNGDQDQFGKTKLTKTAMSN 84
Db 102 PVFIRKDRVTVO-----EREAEAL-----KQKELEQEA-- 131

Qy 85 PFNSSESLSADLFIYKDKPEELKKYLKVRHRTFVVL--INTGNSDDGLKIYKDYD 142
Db 132 --RMAEERRK---YTLKIVEETKKELENKRSIALDALNTDDEND-----EVEE 178

Qy 143 KFKTPSIFVSTKEQIKELFKDKNTEKERNIA-----VYSNNKDNLHLK 189
Db 179 AWKVRCLKRIKRDREDR-EALEKEKAEIERMNLTEERRAEALRANGKVIITNKAVGKYK 237

Qy 190 FISQYLHQASIF--HAVNPGMPLAATPLVD---DTVIGKLTAKINFYSLNLTGLDV 244
Db 238 FLQKYTHRGAFPMDEDEEVYKRDPSAPTELDHFNKTILPKVNVQV--NF-----GRSGR 289

Qy 245 PAFKEGVDLAGAIDEQFTYHYIKNEAIELIRIWNKRNQSKLSALQLSGARD 299
Db 290 TKYTHLVDDTTSFDS-----AWQBSAQNTKFKKQKAAGVRD 327

RESULT 7
US-08-477-451-28
; Sequence 28, Application US/08477451
; Patent No. 592865
; GENERAL INFORMATION:
```



```

; APPLICANT: Covacci, Antonello
; TITLE OF INVENTION: Helicobacter Pylori Cagi Region
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,451
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0335.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2708
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 382 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-477-451-28

```

```

Query Match      6.0%; Score 112.5; DB 1; Length 382;
Best Local Similarity 24.8%; Pred. No. 0.062;
Matches 81; Conservative 40; Mismatches 117; Indels 89; Gaps 18;

QY 27 LVYTKAKIKVNDKANYKILNLTVNNYKQIETLEKONGNGQDQFGKEKTLTKTAMSNFF 86
DB 67 LIFRQKIKELMD-HRAKYSLDNKYKKEKELEKTR-----GK-----ILTKASKAY 115
QY 87 NSSESLKSADLFIYKDKPEELKYLKVHRHTFVVLINTEGNSDD--GLKIYKDDYDKF 144
DB 116 GLEQAL-----KDNP-----LYKLLPNPYAVVLNQETTFKEDKERLSY---YQV 160
QY 145 KPSIF-----FVSTKEQBIKELFKDKGNTKERNIAVYSNNKNL- 186
DB 161 KTSSIFEKTATTDKQAQLQMGVFLDDEQNK--ASRLALSYKQAIIEEYNNISNLL 218
QY 187 ---HLKFTISQYLHQASIFHAVNPYGMPL--AATPLVDDTVIGKLTAKINFYSLNETGL 241
DB 219 SRKELNDIDYQLQLER-----NKFOSKAKDIAQKATNTLIFNSERLAFSMAIDKINEKYL 273
QY 242 DGVPAF-----KEGVDLAGAIDBQFTYHYIKNEAIIELIRIWNKNRQNSKLSALQL 294
DB 274 KYEAFSNLLKNVDDVEL--NTLTKNFT-----NQKL-----SFAQKQKLCCLVL 317
QY 295 SGARDNAVTSATIECLLR-----FVD 315
DB 318 DSF--NFDTSQSKSILKKTNEYNIQVD 342

```

```

RESULT 8
US-09-830-230A-322
; Sequence 322, Application US/09830230A
; Patent No. 6902893
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481US
; CURRENT APPLICATION NUMBER: US/09/830,230A

```

```

; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: PCT/US98/12718
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/057,483
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 322
; LENGTH: 1087
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (541)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (595)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (703)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-830-230A-322

Query Match      5.9%; Score 111; DB 2; Length 1087;
Best Local Similarity 22.6%; Pred. No. 0.43;
Matches 74; Conservative 50; Mismatches 122; Indels 82; Gaps 15;

QY 37 NKDAANYK--ILNLTVNNYKQIETLEKONGNGQDQFGKEKTLTKTAMSN-----FFNS 88
DB 741 NTQEDHYKLGIRFKLKYKHSIESF-----DQTIKLPKHKKALHNGKIALMLNKK 792
QY 89 SEESLSKADLFIYKDKPEELKYLKVHRHTFVVLINTEGNSDDGLKIYKDDYDKFPTPS 148
DB 793 NKKAIESFEKAIQIDKNYGTAYQK-----GIAEKNQDMQQAFAFKNAYNLDKNPN 845
QY 149 IFF---VFST-----KEQIKELFKDKGNTKERNIAVYS-----NNKDNLHKFT- 191
DB 846 YALKAGIVSNMLGNFKQSEEVINFF--NANAKKPEIAIYNLSIAKPENKLEESLETIN 903
QY 192 -----SOYLH-QASIFHAVNPYGMPLAATPLVDDTVIGKLTAKINFYSLNETGL 241
DB 904 KAILDNPEKSEYLYLKASINLKENY-----QNAISLSYSLVIEKNP 944
QY 242 DGVPAFKEGVDLAGAIDBQFTYHYIKNEAIIELIRIWNKNRQNSKLSALQSGARDNA 301
DB 945 ENTSAY-----INLA-----KAYEKGNSQALSTLEKIINKNKK--LALNNGILYKCKKN 994
QY 302 YTSATECLLKFVRDGLIIQYKNLSLT 329
DB 995 YOKAIEFEKAIINSIDIEAKY-NLATTL 1021

```

```

RESULT 9
US-09-830-230A-321
; Sequence 321, Application US/09830230A
; Patent No. 6902893
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481US
; CURRENT APPLICATION NUMBER: US/09/830,230A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: PCT/US98/12718
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/057,483

```



```

; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 321
; LENGTH: 1119
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (573)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (627)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (735)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-830-230A-321

Query Match          5.9%; Score 111; DB 2; Length 1119;
Best Local Similarity 22.6%; Pred. No. 0.45;
Matches 74; Conservative 50; Mismatches 122; Indels 82; Gaps 15;

QY 37 NKDAANYK--ILNLTVNNYEKOIETLEKONGNGQDQFQKTKLLKTAMSN-----FPNS 88
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D6 773 NTQEDHYGLIIRFKLKKYHIESF-----DQTIKLPKPKKALHNGKIALMLNK 824
QY 89 SEESLSADLFYKDKPEELKYLKVHRHTFVVLINTEGDSDDGLKIYKDDYKFKTPS 148
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D6 825 NKKAIESFEKAIQIDKNYGTAYQK-----GIAEERKNGMQQAFKFNAYNLDKPN 877
QY 149 IPF---VFST-----KEQIEKELPKDKGNTKERNIAVYS-----NNKDLHLKEI- 191
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D6 878 YALKAGIVSNLGNFKQSEYLNFF--NANAKPNEIAYNLISIAKFENKLEESLEVIN 935
QY 192 -----SOYLH-QASIFHAVNPGMPLAATPLVDVTGKLTAKINFYSLNETGL 241
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D6 936 KALDLNPEKSEYLYLKASINLKKNY-----QNAISLYSLVIEKNP 976
QY 242 DGVPAFKGVDLAAGADEOFTYHYKNEAIIELIRIWNKNRQNSKLSALQLSGARDNA 301
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D6 977 ENTSAY---INLA-----KAYEKGSKQAISTLEKIINKNK--LALNNLGIYKCKN 1026
QY 302 YTSAECLLKFVDRGLIIQYKNLSLT 329
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D6 1027 YOKAIEIFEKAIINSIDIEAKY-NLATTL 1053

RESULT 10
US-09-248-796A-18397
; Sequence 18397, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18397
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-248-796A-18397

Query Match          5.8%; Score 110; DB 2; Length 1078;
Best Local Similarity 21.6%; Pred. No. 0.52;
Matches 79; Conservative 59; Mismatches 140; Indels 88; Gaps 14;

QY 44 KILNLTVNNYEKOIETLEKONGNGQDQFQKTKLLKTAMSNFPNSSESLKADLFYKD 103
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D6 367 KVLNLT-----KELQT-EKENAESNDKELNEKI--EKLTLNLSTKLETKL-----ED 409
QY 104 KPEELKYLKVHR--HTFVVLINT-----EGNSDDG 133
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D6 410 KEQELAKIQEDHKLLNEKFLVTTNSLSAMKASKKEFETASOKYKELQALKKGNTSST 469
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18397
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-248-796A-18397

```

```

; LENGTH: 1241
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18397

Query Match          5.9%; Score 111; DB 2; Length 1241;
Best Local Similarity 20.9%; Pred. No. 0.53;
Matches 87; Conservative 58; Mismatches 162; Indels 110; Gaps 18;

QY 4 DTISVSLDSRIQASRPNY-----YNPLL--VYKTAKI--KVNKDAANYKIL----- 46
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D6 792 DNTSTTTNTKYIRAPKLKYESMFLIYNDKLLPIYNSATILFKIELVESNKWYKLFGL 851
QY 47 -NLTVNNYEKOIETLEKONGNGQDQF-----GREKTIILKTAMSNFPNSSESL 93
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D6 852 RNLIQIKNTSEDIQKHLNIDANQYFEINDTVDLSTVINEPKTLLNLINFNTLKNLSLI 911
QY 94 KSADLFYKDKPEELKYLKVHRHTFVVLINTEGDSDDGLKIYKDDYKFKTPSIPFV 152
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D6 912 KILD-----QNKSNVTVL-----DMSMDNITINIKDYNNNDKLLIILK 951
QY 153 FSTKEQIEKELPKDKGNTKERNIAVYSNN--KDLHLKFISQYLHQASI----- 200
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D6 952 LPEQATDSIQLMKNGSTTTTDEIDLNNNILEFELILKFLNQYLRESKLNHNRRQIISI 1011
QY 201 -----FHAVNPGMPLAATPLVDVTGKLTAKINFYSLNETGLDGVPAFKGVDLA 254
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D6 1012 IKIIQYLTENPI---LQSTKAINQQLAELKR-----INSTNNNNNAKRRIILKLS 1059
QY 255 GGAIDEOFTYHYKNEAIIELIRIWNKNRQNSK-----LSAL---OLSGARDNA 301
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D6 1060 NGLYKFLVNLNII--SLTHLQVLFVFNSTGSLNKKIQRDKIMINLSLKFDRFSKPN 1118
QY 302 YTSAECLLKFVDRGLI---IOYKNL-----SLTSLPTQLKLESLVNITYN 346
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D6 1119 QDSTNGHLIKSFKDSLINENLKEFNLEIFKKNINELLITKSKQIK---SLNQTEN 1172

RESULT 11
US-09-248-796A-20284
; Sequence 20284, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20284
; LENGTH: 1078
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20284

Query Match          5.8%; Score 110; DB 2; Length 1078;
Best Local Similarity 21.6%; Pred. No. 0.52;
Matches 79; Conservative 59; Mismatches 140; Indels 88; Gaps 14;

QY 44 KILNLTVNNYEKOIETLEKONGNGQDQFQKTKLLKTAMSNFPNSSESLKADLFYKD 103
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D6 367 KVLNLT-----KELQT-EKENAESNDKELNEKI--EKLTLNLSTKLETKL-----ED 409
QY 104 KPEELKYLKVHR--HTFVVLINT-----EGNSDDG 133
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D6 410 KEQELAKIQEDHKLLNEKFLVTTNSLSAMKASKKEFETASOKYKELQALKKGNTSST 469
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20284
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20284

```

Db 470 LQKLEKLDSTEQAKKLEGGDGINNTRDLFLHKLKSEAFTQIKREREFKNLTVELENT 529
 QY 171 EKERNIAVYNNKNLHLK----FISYLHQASIFHAVNPGM--PLAATPLVDDTVIGK 224
 Db 530 KADYELQINLKNSENEFKQKINELUSKKIESUTDNKNAKOLEKJLKDTEENHEHMDK 589
 QY 225 LRTAKINFYSL---LNETGLDGPAPKEGVDLAGGAID----EQTYHYIKNEAIEILIR 277
 Db 590 LRSASVAYNDLKAKSESEETVKA-KEELETLTSKIDNLEKELKEQSKKNELEGGQLQN 648
 QY 278 INWKNRQNSKLSALQSGARDNAVTSIECLLKRFPVDRGLIIQVKNLSLTLSPFPQLK 337
 Db 649 ITDSTNEKFELELDKLSKSNKEISSONSELIOKLEK---TEKOLQAKDEEIDKLKA 704
 QY 338 ELSVNI 343
 Db 705 ETKSNI 710

RESULT 12
 US-08-923-992A-10
 ; Sequence 10, Application US/08923992A
 ; Patent No. 6280738
 ; GENERAL INFORMATION:
 ; APPLICANT: Tai, Joseph Y.
 ; APPLICANT: Blake, Milan S.
 ; TITLE OF INVENTION: No. 6280738-Iga Fc Binding Forms of the Group B
 ; TITLE OF INVENTION: Streptococcal Beta Antigens
 ; NUMBER OF SEQUENCES: 34
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
 ; STREET: 1100 New York Avenue, N.W., Suite 600
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC Compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/923,992A
 ; FILING DATE: 05-SEP-1997
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/024,707
 ; FILING DATE: 06-SEP-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ramond, Robert W.
 ; REGISTRATION NUMBER: 32,893
 ; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 371-2600
 ; TELEFAX: (202) 371-2540
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1164 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-08-923-992A-10

Query Match 5.8%; Score 109; DB 2; Length 1164;
 Best Local Similarity 22.6%; Pred. No. 0.73;
 Matches 83; Conservative 47; Mismatches 134; Indels 104; Gaps 18;
 QY 55 KOIETLEKDNNGQDFGKEKTL-----LKTAMSNF-----FNSSESLKSADLFIYDKK 104
 Db 165 KOHEVEKDKKAKQ-----QKTLKQSDTKVDLSIDKELNHOKSQVETMAEOLGTNEDK 219
 QY 105 PEELKKYLKVRHRTFVVLINTEGDSDDGLKIYKODYDKFKTPTSFVFFSTK---EQBIK 161

Db 220 DSKLKIETDIRKQA-----QQADKEDAEVKKVEELGKLF-----STKAGLDQBIQ 266
 QY 162 ELFKDKGNTKERNIAV---YSNNKNLHLKFI-----SOYLHQASIFH 202
 Db 267 EHV-KETTISENTQKDEHYFPNSLQNLQAQSLSEBELDKATTNEQATQVKNQFLENAQKLK 325
 QY 203 AVNPGMPLAATPLVDDTVIGKLTAKINFYSLNLTGLDGPAPKEGVDLAGGAIDBQF 262
 Db 326 EIQP-----LIKETNVKLYKAMSESL-----QVEKQL 353
 QY 263 TYHYIKN-BAII-----ELIRIWNKNRQNSKLSAL-QLSGARDNAVTSIECLLKRFPVDR 316
 Db 354 KHSQANLELDVAKSKEIVREYEGKLNQSKNLPKLEEBEAHSLKQVWEDFRKFKFNTS 413
 QY 317 GLIIQYKNL--SLTLSPFPQLKELSV---NIT-----YNFSINAVSLVITTDIVD-- 363
 Db 414 EQVTPKRVKRDLAANENNOQKIELTVSPENITVYEGEDLKFRTLTAKSDDSKTTLDFSDLL 473
 QY 364 --YQNSLS 369
 Db 474 TKYNPSVS 481

RESULT 13
 US-09-134-000C-5630
 ; Sequence 5630, Application US/09134000C
 ; Patent No. 6617156
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 032796-032
 ; CURRENT APPLICATION NUMBER: US/09/134,000C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/055,778
 ; PRIOR FILING DATE: 1997-08-15
 ; NUMBER OF SEQ ID NOS: 6812
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 5630
 ; LENGTH: 429
 ; TYPE: PRT
 ; ORGANISM: Enterococcus faecalis
 US-09-134-000C-5630

Query Match 5.7%; Score 107; DB 2; Length 429;
 Best Local Similarity 20.1%; Pred. No. 0.24;
 Matches 71; Conservative 46; Mismatches 128; Indels 108; Gaps 11;
 QY 6 ISVSLDSRIQASRPNYNPLLVYKTAKIKVKNDAANYKIILNTVNNYEKQIETLEKONG 65
 Db 153 ITIALIDSGIDRLHPNLQDN-----NLRKNYVNDIELDEYGHG 191
 QY 66 -----NGDQFGKEKTLTKTAMS-----NFFNSEESLSKAD 97
 Db 192 TVAGVIDTIAPRVNLNSYKVMGTD--GNSINMLKAIVDATNDQVDIINVSLSGSKNME 249
 QY 98 LFIYKDPKEELKKYLKVRHRTFVVLINTEGDSDDGLKIYKDDYDKFKTPTSFVFFSTKE 157
 Db 250 IDDERFTVEAFKAVNYARKNNILIVASAGNESRD-----ISTGNE 290
 QY 158 QEIK---ELFKDKGNTKERNIAVYNNKNLHLKFIISQYLHQASIFHAVNPGMPLAAT 214
 Db 291 KHIFGGLESVITVGATKSGDIADYSNYGSNV-----SIYGPAGYGDNYKIT 338
 QY 215 PLVDDTVIGKLTAKINFYSLNLTGLDGPAPKEGVDLAGGAIDEQTYHYIKNEAIE 274
 Db 339 GOID-----AREMMVTYPTSLVSLGKAADFPDGYTLSPGT-----SLATPEVSAA 385
 QY 275 LIRIWNKNRQNSKLSALQSGARDNAVTSIECLLKRFPVDRGLIIQVKNLSL 327
 Db 386 LAATMSK-NVDNSKSDSNEVLNTLFENA-----DSFIDKNSMLKYKEVRI 428

```

US-09-248-796A-15476

Query Watch      5.6%; Score 105; DB 2; Length 1060;
Best Local Similarity 20.0%; Pred. No. 1.5;
Matches 79; Conservative 72; Mismatches 150; Indels 94; Gaps 18;

QY 10 LLDSRIQASRPYNYPLLVYKTAIKIVNKDAANYKILNTVNNYKQIETLEKDNNGQD 69
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 106 VFNEKVGSPSWM---ILLY-----VPSADKQSTIIKMSHYDKLLKDFSNEGGKELA 157
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 70 QFGKEKTLTKTAMSNFFNSBESLSKSAADIYKDKPEELKKYL-----KVRHHT 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 158 ALFAESTSTASV-----GAQNSKSTGYCFKQKQHEFELNEFLVQIKNLLGFTFNQKYHL 212
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 FVVLINTEGNSDGLHIKYDDYD-----KFKTPSIFVFVSTKEQEIKELFKQKNGT 170
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 213 NLDLIITSGDHENNSLTKYVATYNLAEFLYDMKLFNDCINF--FNRLSEQLNTLVENNPIL 271
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 171 ----EKERNIAVYNSKNDNLHLKFIIS---QYLHQASIFHAVNPYGMPL-----AAT--PL 216
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 272 FIYKVDLPAPKVFNSKFD--FKFPYNNKCOHMDVNSFTNVNLFELKCFIFFRQASTLEML 329
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 217 VDDTVLIGLRTAKINYSILNETGL---DGVPAFKEGVDLAGGAIDQEQYTHYIK----- 268
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 330 VNKNLANSISLAELQISKLRLNLVFLNDLQVQNEQALIT-----EFEYSIIEYFLNL 383
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 269 ---NEAIIELIRTW-----NKN-----RQNSKLISALQLSGARDNAVYTSATECLLKR 312
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 384 DIVNKLIQKATKIYEAEPANNNSYOLKRLFESRGELKLLQRSLSIKLARENSIEI----- 439
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 313 FVDRGLHIIQYKNLSL-----TLSPTPOLKLEL 339
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 440 ---KGLDQVFDVSLDEEGPKQSKDSTEQVKLDL 471
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Search completed: January 24, 2006, 20:03:01
Job time : 23.8097 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:57:15 ; Search time 65.0296 Seconds
(without alignments)
2377.329 Million cell updates/sec

Title: US-10-688-058-20
Perfect score: 1889
Sequence: 1 MPQDTISVSLDSRIQASRP.....AVSLVITQDIDVYQNSLSA 370

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_Main:
1: /cgm2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgm2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgm2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgm2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgm2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgm2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1889	100.0	370	5	US-10-688-058-20 Sequence 20, Appl
2	1837	97.2	370	5	US-10-688-058-64 Sequence 64, Appl
3	788	41.7	160	5	US-10-688-058-62 Sequence 62, Appl
4	329	17.4	65	5	US-10-688-058-18 Sequence 18, Appl
5	137.5	7.3	1633	4	US-10-282-122A-70437 Sequence 70437, A
6	130.5	6.9	1742	4	US-10-615-383-4 Sequence 4, Appl
7	130.5	6.9	1742	4	US-10-690-184-4 Sequence 4, Appl
8	130.5	6.9	1742	4	US-10-689-082-4 Sequence 4, Appl
9	129.5	6.9	1877	5	US-10-732-923-22524 Sequence 22524, A
10	129.5	6.9	1877	5	US-10-732-923-22547 Sequence 22547, A
11	120.5	6.4	575	4	US-10-724-972A-7483 Sequence 7483, Ap
12	120.5	6.4	598	4	US-10-282-122A-52018 Sequence 52018, A
13	117.8	6.3	1178	4	US-10-282-122A-72168 Sequence 72168, A
14	117.5	6.3	1331	4	US-10-282-122A-63427 Sequence 63427, A
15	117.5	6.2	5005	4	US-10-282-122A-76871 Sequence 76871, A
16	117	6.2	806	5	US-10-732-923-23131 Sequence 23131, A
17	117	6.2	1156	4	US-10-369-493-43 Sequence 43, Appl
18	116.5	6.2	1191	4	US-10-282-122A-52048 Sequence 52048, A
19	116.5	6.2	6761	5	US-10-732-923-15035 Sequence 15035, A
20	115	6.1	590	4	US-10-282-122A-52497 Sequence 52497, A
21	113.5	6.0	596	4	US-10-282-122A-52435 Sequence 52435, A
22	113	6.0	1805	3	US-09-820-843A-73 Sequence 73, Appl
23	113	6.0	1805	4	US-10-282-122A-63513 Sequence 63513, A
24	113	6.0	1805	5	US-10-732-923-3303 Sequence 3303, Ap
25	112.5	6.0	1265	4	US-10-032-585-7168 Sequence 7168, Ap
26	112	5.9	1394	4	US-10-437-963-183281 Sequence 183281, A
27	112	5.9	1881	4	US-10-032-585-7646 Sequence 7646, Ap

ALIGNMENTS

RESULT 1

US-10-688-058-20
; Sequence 20, Application US/10688058
; Publication No. US20050058661A1
; GENERAL INFORMATION:
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: HALE, KATHERINE S.
; APPLICANT: JOHNSTON, STEPHEN A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF THE GENUS
; TITLE OF INVENTION: BORRELIA
; FILE REFERENCE: MCR0:00305
; CURRENT APPLICATION NUMBER: US/10/688,058
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
; US-10-688-058-20

Query Match	100.0%	Score 1889;	DB 5;	Length 370;
Best Local Similarity	100.0%;	Pred. No. 5e-144;		
Matches 370;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MPQDTISVSLDSRIQASRPNYNPLLVYKTAKIKVNDAAANYKILNLTNNYKQIETL	60	
Db	1	MPQDTISVSLDSRIQASRPNYNPLLVYKTAKIKVNDAAANYKILNLTNNYKQIETL	60	
Qy	61	EKONGNGODQFGKTKLLKTAMSNFNSSESLSKADLIYKDKPEELKKYLKVRHFTV	120	
Db	61	EKONGNGODQFGKTKLLKTAMSNFNSSESLSKADLIYKDKPEELKKYLKVRHFTV	120	
Qy	121	VLINTEGNSDDGLKIYKDDYDKETPSIFPVFSTKEQEIKELPDKGKTEKERNIAVYS	180	
Db	121	VLINTEGNSDDGLKIYKDDYDKETPSIFPVFSTKEQEIKELPDKGKTEKERNIAVYS	180	
Qy	181	NNKDNHLKFTISQYLHQASIFHAVNPGMPLAATPLVDDTVIGKLTAKINFYSLNBTG	240	
Db	181	NNKDNHLKFTISQYLHQASIFHAVNPGMPLAATPLVDDTVIGKLTAKINFYSLNBTG	240	
Qy	241	LDGVPAPKEGVDLAGGAIDQFTYHIKNEAIIELIRIWNKNRNSKLSALQLSGARDN	300	
Db	241	LDGVPAPKEGVDLAGGAIDQFTYHIKNEAIIELIRIWNKNRNSKLSALQLSGARDN	300	
Qy	301	AYTSAIECLLKRFDVRGLIIQYKNLSLTLSPTQPKLELSVNTITVNFISINAVSLVITQD	360	
Db	301	AYTSAIECLLKRFDVRGLIIQYKNLSLTLSPTQPKLELSVNTITVNFISINAVSLVITQD	360	

```
Qy 361 IVDYQNSLSA 370
|||||
Db 361 IVDYQNSLSA 370

RESULT 2
US-10-688-058-64
; Sequence 64, Application US/10688058
; Publication No. US20050058661A1
; GENERAL INFORMATION:
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: JOHNSTON, STEPHEN A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF THE GENUS
; TITLE OF INVENTION: BORRELIA
; FILE REFERENCE: MCRO:003US
; CURRENT APPLICATION NUMBER: US/10/688,058
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-688-058-64

Query Match 97.2%; Score 1837; DB 5; Length 370;
Best Local Similarity 97.0%; Pred. No. 7.9e-140;
Matches 359; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MPQDTISVSLDSRIQASRPYNNPLLVYKTAKIKVKNDAANYKILNLTNNYKQIETL 60
|||||
Db 1 MPQDTISVSLDSRIQASRPYNNPLLVYKTAKIKVKNDAASYKILNLTNNYKQIETL 60
|||||

Qy 61 EKXNGNGDQFQKCKTLKTAMSNFFNSSEESLKSADLFYKDKPEELKYLKVRHTFV 120
|||||
Db 61 EKXNGNGDQFQKCKTLKTAMSNFFNSSEESLKSADLFYKDKPEELKYLKVRHTFV 120
|||||

Qy 121 VLINTEGNSDGLKIYDDYDKFTPTSIFFVSTKEQBIKELFKDKGTEKERNIAVYS 180
|||||
Db 121 VLINTEGASDGLKIYDDYKNFKPSTFFVSTKEQBIKELFKDKGTEKERNIAVYS 180
|||||

Qy 181 NKNKDLHLKFIQYLHQASIFHAVNPGMPLAATPLVDVTIGKLTAKINFYSLNETG 240
|||||
Db 181 NKNKDLHLKFIQYLHQASIFHAVNPGMPLAATPLVDVTIGKLTAKINFYSLNETG 240
|||||

Qy 241 LDGVPAPKEGVDLAGGAIDQFTYHYIKNEAIIELIRIWNKNNRQNSKLSALQLSGARDN 300
|||||
Db 241 LDGVPAPKEGVDLAGGAIDQFTYHYIKNEAIIELIRIWNKNNRQNSKLSALQLSGARDN 300
|||||

Qy 301 AYTSAEICLLKRFVDRGLIIQVKNLSLTSPFPQKLELSVNITNFNSINAVSLVITQD 360
|||||
Db 301 AYTSAEICLLKRFVDRGLIIQVKNLSLTSPFPQKLELSVNITNFNSINAVSLVITQD 360
|||||

Qy 361 IVDYQNSLSA 370
|||||
Db 361 IVDYQNSLSA 370
|||||

RESULT 3
US-10-688-058-62
; Sequence 62, Application US/10688058
; Publication No. US20050058661A1
; GENERAL INFORMATION:
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: JOHNSTON, STEPHEN A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF THE GENUS
; TITLE OF INVENTION: BORRELIA
; FILE REFERENCE: MCRO:003US
; CURRENT APPLICATION NUMBER: US/10/688,058

Query Match 97.2%; Score 1837; DB 5; Length 370;
Best Local Similarity 97.0%; Pred. No. 7.9e-140;
Matches 359; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MPQDTISVSLDSRIQASRPYNNPLLVYKTAKIKVKNDAANYKILNLTNNYKQIETL 60
|||||
Db 1 MPQDTISVSLDSRIQASRPYNNPLLVYKTAKIKVKNDAASYKILNLTNNYKQIETL 60
|||||

Qy 61 EKXNGNGDQFQKCKTLKTAMSNFFNSSEESLKSADLFYKDKPEELKYLKVRHTFV 120
|||||
Db 61 EKXNGNGDQFQKCKTLKTAMSNFFNSSEESLKSADLFYKDKPEELKYLKVRHTFV 120
|||||

Qy 121 VLINTEGNSDGLKIYDDYDKFTPTSIFFVSTKEQBIKELFKDKGTEKERNIAVYS 180
|||||
Db 121 VLINTEGASDGLKIYDDYKNFKPSTFFVSTKEQBIKELFKDKGTEKERNIAVYS 180
|||||

Qy 181 NKNKDLHLKFIQYLHQASIFHAVNPGMPLAATPLVDVTIGKLTAKINFYSLNETG 240
|||||
Db 181 NKNKDLHLKFIQYLHQASIFHAVNPGMPLAATPLVDVTIGKLTAKINFYSLNETG 240
|||||

Qy 241 LDGVPAPKEGVDLAGGAIDQFTYHYIKNEAIIELIRIWNKNNRQNSKLSALQLSGARDN 300
|||||
Db 241 LDGVPAPKEGVDLAGGAIDQFTYHYIKNEAIIELIRIWNKNNRQNSKLSALQLSGARDN 300
|||||

Qy 301 AYTSAEICLLKRFVDRGLIIQVKNLSLTSPFPQKLELSVNITNFNSINAVSLVITQD 360
|||||
Db 301 AYTSAEICLLKRFVDRGLIIQVKNLSLTSPFPQKLELSVNITNFNSINAVSLVITQD 360
|||||

Qy 361 IVDYQNSLSA 370
|||||
Db 361 IVDYQNSLSA 370
|||||

RESULT 4
US-10-688-058-18
; Sequence 18, Application US/10688058
; Publication No. US20050058661A1
; GENERAL INFORMATION:
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: HALE, KATHERINE S.
; APPLICANT: JOHNSTON, STEPHEN A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF THE GENUS
; TITLE OF INVENTION: BORRELIA
; FILE REFERENCE: MCRO:003US
; CURRENT APPLICATION NUMBER: US/10/688,058
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-688-058-18

Query Match 17.4%; Score 329; DB 5; Length 65;
Best Local Similarity 100.0%; Pred. No. 4.3e-19;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 211 LAATPLVDDTVIGKLTAKINFYSLNETGLDGVPAKFGVDLAGGAIDQFTYHYIKNE 270
|||||
Db 1 LAATPLVDDTVIGKLTAKINFYSLNETGLDGVPAKFGVDLAGGAIDQFTYHYIKNE 270
|||||

Qy 271 AIIELIRIWNKNNRQNSKLSALQLSGARDNAYTSAIECLLKRFVDRGLIIQVKNLSLTLS 330
|||||
Db 61 AIIELIRIWNKNNRQNSKLSALQLSGARDNAYTSAIECLLKRFVDRGLIIQVKNLSLTLS 120
|||||

Qy 331 PTPOLKLELSVNITNFNSINAVSLVITTDIVDYQNSLSA 370
|||||
Db 121 PTPOLKLELSVNITNFNSINAVSLVITTDIVDYQNSLSA 160
|||||

RESULT 5
US-10-282-122A-70437
; Sequence 70437, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
```

APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELTRA 03AA
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patent in version 3.1
SEQ ID NO 70437
LENGTH: 1633
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-10-282-122A-70437

Query Match 7.3%; Score 137.5; DB 4; Length 1633;
Best Local Similarity 22.6%; Pred. No. 0.11;
Matches 97; Conservative 58; Mismatches 168; Indels 107; Gaps 21;
Qy 18 SRPNYNPLLVYTKAKIVNK-DAANYKILNLTNNYKQIETLEKNGQDQFGKEKT 76
Db 113 SKPRANEAVVTNESTKPKTTEAPTVEESIAETPKTSTTQODSTKKNPSLKNLNSST 172
Qy 77 LLKTAMSNFFNSSESLKADLFYK-----DKPEELKYLKVRHHTFVVLINTEGNS 130
Db 173 TSKE-----KTDEHSTKQAMSTKNSLNDTNDSTQSEK-----TSSQANDSTNQ 220
Qy 131 D-----DGLKIYKDDYD-----KFKTPSIPFFVSTKEQE-----IKE 162
Db 221 SAPSKQLDSKSEQKVKTKFNDEPTQDVEYTTTKLTPSISTDSSVNDKQDYTSASVAS 280
Qy 163 LFKDKGNTKERNIAVYNNKNDLHLKFPIS-OYLHOASIFHAV-----NP-YGM--PLAA 213
Db 281 LGVDSNETE-----AITNAVRDNLKTAAREQINEAIIAALKKDFNSPDYGVDTPLAL 335
Qy 214 TPLVDDTVIGKRTAKINFYSLNETGL-----DGVPAFKEGVDL-----AGGAI----- 258
Db 336 NTSQKTSFKSASPRMNLMSLAABPNSGKNVNDKVKITNPTLSLNKNNHANNVWPTS 395
Qy 259 DEQFTY--HYIKNEAI-----TELIRIWNKNNRQNSKLSALQSGAR 298
Db 396 NEQFNLKANYELDDSIKEGDTFTIKYQYIRPGGLELPAI---KTQLRSKDGSIANGVY 452
Qy 299 DNAYTSAIECLLKRFVDRGLIIQYKNL--SLTSPTPOLKLELSVNTYFNINAVSLVI 356
Db 453 DKT-TNTTYYTFTNYVD-----QYQNTIGSPDLIATPKRETAIKONQYPMVEITIANEV 506
Qy 357 TTQDIDVYQN 366
Db 507 KDFIVDYG 516

RESULT 6

US-10-615-383-4
Sequence 4, Application US/10615383
Publication No. US20040038327A1
GENERAL INFORMATION:
APPLICANT: POSTER, Timothy
TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCCI
FILE REFERENCE: P06335US03/BAS
CURRENT APPLICATION NUMBER: US/10/615,383
CURRENT FILING DATE: 2003-07-09
PRIOR APPLICATION NUMBER: 09/386,962
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/098,443
PRIOR FILING DATE: 1998-08-31
PRIOR APPLICATION NUMBER: 60/117,119
PRIOR FILING DATE: 1999-01-25
NUMBER OF SEQ ID NOS: 39
SOFTWARE: Patent in version 3.1
SEQ ID NO 4
LENGTH: 1742
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-10-615-383-4

Query Match 6.9%; Score 130.5; DB 4; Length 1742;
Best Local Similarity 22.1%; Pred. No. 0.44;
Matches 95; Conservative 59; Mismatches 169; Indels 107; Gaps 21;
Qy 18 SRPNYNPLLVYTKAKIVNK-DAANYKILNLTNNYKQIETLEKNGQDQFGKEKT 76
Db 122 SKPRANEAVVTNESTKPKTTEAPTVEESIAETPKTSTTQODSTKKNPSLKNLNSST 181
Qy 77 LLKTAMSNFFNSSESLKADLFYK-----DKPEELKYLKVRHHTFVVLINTEGNS 130
Db 182 TSKE-----KTDEHSTKQAMSTKNSLNDTNDSTQSEK-----TSSQANDSTNQ 229
Qy 131 D-----DGLKIYKDDYD-----KFKTPSIPFFVSTKEQE-----IKE 162
Db 230 SAPSKQLDSKSEQKVKTKFNDEPTQDVEHHTTKLTPSVSTDSSVNDKQDYTSASVAS 289
Qy 163 LFKDKGNTKERNIAVYNNKNDLHLKFPIS-OYLHOASIFHAV-----NP-YGM--PLAA 213
Db 290 LGVDSNETE-----AITNAVRDNLKTAAREQINEAIIAALKKDFNSPDYGVDTPLAL 344
Qy 214 TPLVDDTVIGKRTAKINFYSLNETGL-----DGVPAFKEGVDL-----AGGAI----- 258
Db 345 NRSQKNSPKSASPRMNLMSLAABPNSGKNVNDKVKITNPTLSLNKNNHANNVWPTS 404
Qy 259 DEQFTY--HYIKNEAI-----TELIRIWNKNNRQNSKLSALQSGAR 298
Db 405 NEQFNLKANYELDDSIKEGDTFTIKYQYIRPGGLELPAI---KTQLRSKDGSIANGVY 461
Qy 299 DNAYTSAIECLLKRFVDRGLIIQYKNL--SLTSPTPOLKLELSVNTYFNINAVSLVI 356
Db 462 DKT-TNTTYYTFTNYVD-----QYQNTIGSPDLIATPKRETAIKONQYPMVEITIANEV 515
Qy 357 TTQDIDVYQN 366
Db 516 KDFIVDYG 525

RESULT 7

US-10-690-184-4
Sequence 4, Application US/10690184
Publication No. US20040141997A1
GENERAL INFORMATION:
APPLICANT: POSTER, Timothy
TITLE OF INVENTION: METHODS FOR TREATING OR PREVENTING INFECTIONS FROM COAGULASE-NEGATIVE STAPHYLOCOCCI
FILE REFERENCE: P06335US05/BAS
CURRENT APPLICATION NUMBER: US/10/690,184

```
; CURRENT FILING DATE: 2003-10-21
; PRIOR APPLICATION NUMBER: 09/386,962
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/098,443
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/117,119
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1742
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-690-184-4

Query Match          6.9%; Score 130.5; DB 4; Length 1742;
Best Local Similarity 22.1%; Pred. No. 0.44;
Matches 95; Conservative 59; Mismatches 169; Indels 107; Gaps 21;

Qy 18 SRPNYNNPLLVYKTAIKVVK-DAANYKILNLTVNNYEKQIETLEKONGNGQDQFGKEKT 76
Db 122 SKPKANEAVVNESTKPKTTEAPTNEESIAETPKTSTTQDSTKKNPSSLKDNLSNST 181
Qy 77 LLKTAMSNFFNSSELSKADLFYK-----DKPEELKKYLKVRHRTFVVLINTEGDS 130
Db 182 TSKE-----KTDEHSTKQAQMSNLSNDTNDSTQSEK-----TSSQANNDSTNQ 229
Qy 131 D-----DGLKIYKDDYD-----KFKTPSIFVFSYKEQE-----IKE 162
Db 230 SAPSKQLDSKPSKQVYKTKFNDEPTQDVEHTTTKLKTPSVSTSSVNDKQDYTRSAVAS 289
Qy 163 LFKDKGNTKERNIAVYNNKDNHLKPFIS-QYLHOASIFHAV-----NP-YGM--PLAA 213
Db 290 LGVDSNETE-----AITNAVRNDLDKAAREQINEALIAEALKKDFSNPDYGVDTPLAL 344
Qy 214 TPLVDDTVIGKLTAKINFYSLNLTGL-----DGVPAFKEGVDL-----AGGAI----- 258
Db 345 NRSQSKNSPHKSPRNLMSLAAPNSGKNVNDVKITNPTLSLNKSNHANNVWPTS 404
Qy 259 DRQFTY--HYIKNEAI-----IELIRWNKNRQNSKLSALQLSGAR 298
Db 405 NEQFNLKANYELDDSIKEGDFTTIKYQYIRPGGLELPAI-----KTQLRSKDGSI VANGVY 461
Qy 299 DNAYTSAIECLLKRFVDRGLIIQYKNL--SLTSPQLKLELSVNITNFNSINAVSLVI 356
Db 462 DKT-TNTTYYTFTNYVD-----QYQNTGSPDLIATPKRETAIKONQNPMEVTIANEVV 515
Qy 357 TTQDIVDYQN 366
Db 516 KKDFIVDYN 525

RESULT 8
US-10-689-082-4
; Sequence 4, Application US/10689082
; Publication No. US20040142348A1
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy
; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLO
; FILE REFERENCE: P06335US04/BAS
; CURRENT APPLICATION NUMBER: US/10/689,082
; CURRENT FILING DATE: 2003-10-21
; PRIOR APPLICATION NUMBER: 09/386,962
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/098,443
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/117,119
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1742
; TYPE: PRT
```

```
; ORGANISM: Staphylococcus epidermidis
US-10-689-082-4

Query Match          6.9%; Score 130.5; DB 4; Length 1742;
Best Local Similarity 22.1%; Pred. No. 0.44;
Matches 95; Conservative 59; Mismatches 169; Indels 107; Gaps 21;

Qy 18 SRPNYNNPLLVYKTAIKVVK-DAANYKILNLTVNNYEKQIETLEKONGNGQDQFGKEKT 76
Db 122 SKPKANEAVVNESTKPKTTEAPTNEESIAETPKTSTTQDSTKKNPSSLKDNLSNST 181
Qy 77 LLKTAMSNFFNSSELSKADLFYK-----DKPEELKKYLKVRHRTFVVLINTEGDS 130
Db 182 TSKE-----KTDEHSTKQAQMSNLSNDTNDSTQSEK-----TSSQANNDSTNQ 229
Qy 131 D-----DGLKIYKDDYD-----KFKTPSIFVFSYKEQE-----IKE 162
Db 230 SAPSKQLDSKPSKQVYKTKFNDEPTQDVEHTTTKLKTPSVSTSSVNDKQDYTRSAVAS 289
Qy 163 LFKDKGNTKERNIAVYNNKDNHLKPFIS-QYLHOASIFHAV-----NP-YGM--PLAA 213
Db 290 LGVDSNETE-----AITNAVRNDLDKAAREQINEALIAEALKKDFSNPDYGVDTPLAL 344
Qy 214 TPLVDDTVIGKLTAKINFYSLNLTGL-----DGVPAFKEGVDL-----AGGAI----- 258
Db 345 NRSQSKNSPHKSPRNLMSLAAPNSGKNVNDVKITNPTLSLNKSNHANNVWPTS 404
Qy 259 DRQFTY--HYIKNEAI-----IELIRWNKNRQNSKLSALQLSGAR 298
Db 405 NEQFNLKANYELDDSIKEGDFTTIKYQYIRPGGLELPAI-----KTQLRSKDGSI VANGVY 461
Qy 299 DNAYTSAIECLLKRFVDRGLIIQYKNL--SLTSPQLKLELSVNITNFNSINAVSLVI 356
Db 462 DKT-TNTTYYTFTNYVD-----QYQNTGSPDLIATPKRETAIKONQNPMEVTIANEVV 515
Qy 357 TTQDIVDYQN 366
Db 516 KKDFIVDYN 525

RESULT 9
US-10-732-923-22524
; Sequence 22524, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 22524
; LENGTH: 1877
; TYPE: PRT
; ORGANISM: Plasmodium yoelii
US-10-732-923-22524

Query Match          6.9%; Score 129.5; DB 5; Length 1877;
Best Local Similarity 21.4%; Pred. No. 0.59;
Matches 83; Conservative 51; Mismatches 122; Indels 131; Gaps 19;

Qy 9 SLLDSRTOASRP-----NYYNPLLVTAKIKVKNDAANYKILNLTVNNYEKQIETLEK 62
Db 1238 TLKDSRI---RPIILTGDNAYNCLYV-----GNKIGLFNNNTTYE-SFYSLVN 1281
Qy 63 DNGNGQDQFGKEKTLTKTAMSNFNSSELSK-----ADLFYKCKPEELKKYLKVRH 116
Db 1282 NSNNIDHSNKKCN-----TNFSNNNSFNLSDDTKFMPFENLFKYSSSKPKFNKNVNO 1335
Qy 117 HTFVVLINTEGDSDDGLKIYKDDYDKFTPSIFVFSYKEQEIKELFKDKGNTKERNI 176
```



```

Db      1336  -----IYDNNNDNDD-----NDND-----DDNDNDNDNN 1362
Qy      177  AVYSNKK-----DNLHLKFIQYLHQASIFHAVNPYGMPLA-----AT 214
Db      1363  NSNSNNNDDDDNNNGVAKKQKQNSKYNTINTINQYSGNILDIDASKIEBYLLESYKNN 1422
Qy      215  PLVDDT-----VIGKLTAKINPYSLLNETGLDV-----PAKEGYDLAGGAIDEQ 261
Db      1423  FVNDTTYDQNTIVGYGYLLNELFVNIHNDKNKINTIVLYKDIYKE-IILTGEA---- 1477
Qy      262  FTYHYKNEAIEIIRIWNNKRNQSKLSALQSLGARDNAVTSIAECLLKRFVDRGLIIQ 321
Db      1478  --NYIED-----HIFKIQNEFENETS-----NLHTNNNNNT-----FKQNPPEQ 1517
Qy      322  YKNSLTL-----SPTPOLKLELSVNIT 344
Db      1518  YKNFLKVRIFSRLTNNKMEIHKDFT 1544

RESULT 10
US-10-732-923-22547
; Sequence 22547, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 22547
; LENGTH: 1879
; TYPE: RPT
; ORGANISM: Plasmodium yoelii yoelii
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1879)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-732-923-22547

```

RESULT 11

US-10-724-972A-7483

US-10-724-972A-7483

Sequence 7483, Application US/10724972A

Publication No. US20040147734A1

GENERAL INFORMATION:

APPLICANT: Doucette-Stamm, Lynn

APPLICANT: Bush, David

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: PATH03-16

CURRENT APPLICATION NUMBER: US/10/724,972A

CURRENT FILING DATE: 2003-12-01

PRIOR APPLICATION NUMBER: 09/450,969

PRIOR FILING DATE: 1999-11-29

PRIOR APPLICATION NUMBER: 09/134,001

PRIOR FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: 60/064,964

PRIOR FILING DATE: 1997-11-08

PRIOR APPLICATION NUMBER: 60/055,779

PRIOR FILING DATE: 1997-08-14

NUMBER OF SEQ ID NOS: 7544

SEQ ID NO 7483

LENGTH: 575

TYPE: PRT

ORGANISM: S.epidermidis

US-10-724-972A-7483

Query Match 6.4%; Score 120.5; DB 4; Length 575;

Best Local Similarity 20.6%; Pred. No. 0.61;

Matches 84; Conservative 68; Mismatches 131; Indels 125; Gaps 22;

Qy 29 YKTAKIKVNKDA-ANYKILMLTVNNYEKQLETKKNGN-GODQFGKTKLLKTAMS--- 83

Db 69 YKAWKKNFDEQVQKGAWSNIIAPIVQKRVNVEIKDONGNVRDNGNPKTERKPVITGYR 128

Qy 84 --NFFNSSESLSKADLFYVKDPEELKYLKVRHTFVVLINTEGNSDDGLKIYKDDY 141

Db 129 AHNVFVDAD--TKGKLITAKD-----LIKTEFENSND---YKDL 164

Qy 142 DKK-----TPSI---FPV-----FSTKEQEI-----KELFKDK 167

Db 165 NEFQVINEELTPSVEKHFLDPTLSNGAKGYSPKSDIEIVISDDLSYDMRFKTLIHEY 224

Qy 168 GNTKERNIAVSNKNLHLKFIQYLHOASIFHAVNPYCMPLAATPL-----V 217

Db 225 AHSQHLNN--DIGKQISEHSRSLKEATEASSAVVANYGLDTSYSLGVLSCWGHNIS 282

Qy 218 DDTVIGKLRPAKINFYSLANET--GLDGVPAFKGVDLGAGIDEQTFYHYIKNEAIEL 275

Db 283 DDELKAHIK---NIHSFAKTTEEINSLPEFSQYID---KLESEM-----NKDVVKD 329

Qy 276 IRTWKNKRON--SKLSALO-----LSGARDNATSAIECLLKRFVDRGLIIQYKNL 325

Db 330 LSTMTDLNKLNGFKDITIIKGNLVNDYGLNEISENSYEND-----DFKVNIDYKGF 380

Qy 326 SLTISPPLQKLEI-----SVNITNFS-----INAVSLVITTD 360

Db 381 N-TNNSODQAKIELINKHDDSLARDNYFTQTYNRNVINNTTTFVQDD 427

RESULT 12

US-10-282-122A-54216

US-10-282-122A-54216

Sequence 54216, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari

APPLICANT: Zyskind, Judith

```

; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54216
; LENGTH: 598
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
US-10-282-122A-54216

Query Match          6.4%; Score 120.5; DB 4; Length 598;
Best Local Similarity 22.5%; Pred. No. 0.65;
Matches 97; Conservative 69; Mismatches 133; Indels 133; Gaps 25;

QY 10 LLDRIQASRPNYPLLVYKTAKIKVNDAAANYKILNLTVN---NYEKQIETL----- 60
Db 171 LLD--LKATFPN-LKADFFKAVDNVFKIINNKISVSKNLNHNLENTHHTTTHSMQ 227
QY 61 ----EKDNGNQDQFGKEKTLTKTAMSNFFN--SSEESLSKADLFYKPKPEELKKYLKVH 115
Db 228 KTNKDSG-----SLLSQTLKNLDSILSSKES-----KHEKNDKVK$-KIE 267
QY 116 RHTFVVLINTEGNSDDGLKIYKDDYDKFTPSIFPVFSTKEQ-----EIKELPKD---K 167
Db 268 EDT-----TDAKNTLNKIK--NDEFKAKNLTEL-NIKDKNQDNLNKESKDLNKFNKE 318
QY 168 GNTKERNIAVYNNKND-NLHLKFTISOYLHQASIFHAVNPYGMPLAATPLVDDT--VIGK 224
Db 319 LNKNGEKNLNGENTQDQKNKLNNDONLNDNLN-------KEIVKDTQKLKSVN 367
QY 225 LRTAKINFYSLNLTGLDGPAPFKEGVDLAGGAIDEQFTYHYIKNEAIELIRIWNK--- 281
Db 368 LQAKDFN---LNKEPKNN--NKENKDIKQNFQDKLNFENLNKTQVVQ-----NKNN 415
QY 282 ----NNRQNSK-----LSAQLSGARDNAYTSATEC 308
Db 416 ANFNNNNTNKKETFTQEQTKTHSENVVDKNSLDELNSLVKDLNKNVQTNNARNITPKETLQY 475
QY 309 L---LKRFRVDRGLIIOYK---NLSLTLSPTQLKLELSV-----NITNFSINAVSLVI 356
Db 476 FQSDLKEAVD-----QYKAPITKLSITLNNLGEVEVTLIQRGNNLHFNFSNANML 530
QY 357 TTQDIDVYQNSL 368
|: : : : :|

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72168
; LENGTH: 1178
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-282-122A-72168

Query Match          6.3%; Score 119; DB 4; Length 1178;
Best Local Similarity 22.6%; Pred. No. 2.2;
Matches 67; Conservative 50; Mismatches 118; Indels 62; Gaps 16;

QY 14 RIQASRPNY-----NPLLVTYKTAKIKVNDAAANYKILNLT--VNNYKQIETL-- 60
Db 250 KVKONLAAYYQERDLLETNKTQTLKEKRHLQSRQMDQKQADLLEMTLRLISDYERQIERIHL 309
QY 61 -----EKDNGNQDQFGKEKTLTKTAMSN---FFNSSEESLSKADLFYKPKPEELK 109
Db 310 EVSQTEKTKQTSQKLEQSEQKASLKTFLAQKEVDLEQLEEDLKKQKQEI-KNVETELUS 368
QY 110 KY-----LKVHRHTFVVLINTEGNSDDGLKIYKDDYDKFTPSIFPVFSTKEQEIKE 162
Db 369 RFATDPDHIIESLREDFVRLMQKEADTSNQ-LAVLKAEMDSRKQES-----ESKTAETKQ 422
QY 163 LFKD--KNTKERNIAVYNNKONLHLKFTISOYLHQASIFHAVN-PYG-MPLAATPLVD 218
Db 423 VQADLEKAKREQRESANFETAKTKVQ-ELLKDYQKTAQLVQNLEAAVYEQEOEAMFQLLD 481
```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2006, 19:59:25 ; Search time 5.92175 Seconds
(without alignments)
633.203 Million cell updates/sec

Title: US-10-688-058-20

Perfect score: 1889

Sequence: 1 MPQDTISVSLDSRIQSRP.....AVSLVTTQDIDVYQNSLSA 370

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 70609 seqs, 10134256 residues

Total number of hits satisfying chosen parameters: 70609

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA New:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08 NEW PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US06 NEW PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US07 NEW PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US09 NEW PUB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US10 NEW PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US11 NEW PUB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US12 NEW PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US13 NEW PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	132.5	7.0	US-10-793-626-1780	Sequence 1780, Ap
2	107	5.7	US-10-793-626-38	Sequence 38, Appl
3	103.5	5.5	US-10-467-657-2208	Sequence 2208, Ap
4	102.5	5.4	US-11-051-453-42	Sequence 42, Appl
5	102	5.4	US-11-052-554A-91	Sequence 91, Appl
6	100.5	5.3	US-11-051-453-41	Sequence 41, Appl
7	98	5.2	US-11-055-822-1066	Sequence 1066, Ap
8	96	5.1	US-10-512-109-27	Sequence 27, Appl
9	95.5	5.1	US-10-793-626-440	Sequence 440, App
10	94	5.0	US-11-069-642-20	Sequence 20, Appl
11	94	5.0	US-10-485-517-141	Sequence 141, App
12	94	5.0	US-11-013-759-11	Sequence 11, Appl
13	93.5	4.9	US-11-155-288-20	Sequence 20, Appl
14	93	4.9	US-11-055-822-1070	Sequence 1070, Ap
15	93	4.9	US-10-513-786-2	Sequence 2, Appl1
16	93	4.9	US-10-513-786-1	Sequence 1, Appl1
17	92.5	4.9	US-10-873-528-156	Sequence 156, App
18	92.5	4.9	US-11-196-475-76	Sequence 76, Appl
19	92.5	4.9	US-10-793-626-1230	Sequence 1230, Ap
20	92.5	4.9	US-10-793-626-2594	Sequence 2594, App
21	92	4.9	US-10-793-626-786	Sequence 786, App
22	92	4.9	US-11-074-176-62	Sequence 62, Appl
23	91.5	4.8	US-11-196-475-66	Sequence 66, Appl
24	91	4.8	US-11-052-554A-69	Sequence 69, Appl
25	91	4.8	US-10-485-517-145	Sequence 145, App

26	90.5	4.8	212	6	US-10-793-626-1128	Sequence 1128, Ap
27	90.5	4.8	622	7	US-11-070-080-22	Sequence 22, Appl
28	90	4.8	302	6	US-10-793-626-2980	Sequence 2980, Ap
29	90	4.8	522	6	US-10-793-626-604	Sequence 604, App
30	89.5	4.7	438	7	US-11-196-475-168	Sequence 168, App
31	89	4.7	383	6	US-10-517-939-334	Sequence 334, App
32	89	4.7	2804	7	US-11-120-925-3	Sequence 3, Appli
33	88.5	4.7	383	6	US-10-793-626-386	Sequence 386, App
34	88.5	4.7	433	7	US-11-196-475-180	Sequence 180, App
35	88	4.7	307	6	US-10-793-626-522	Sequence 522, App
36	88	4.7	376	6	US-10-517-939-102	Sequence 102, App
37	88	4.7	395	6	US-10-793-626-668	Sequence 668, App
38	88	4.7	1167	7	US-11-052-554A-121	Sequence 121, App
39	88	4.7	2504	6	US-10-647-956A-8	Sequence 8, Appli
40	87.5	4.6	2515	7	US-11-113-424-53	Sequence 53, Appl
41	87	4.6	832	6	US-10-512-109-29	Sequence 29, Appl
42	86.5	4.6	1147	6	US-10-615-668-5	Sequence 5, Appli
43	86	4.6	615	7	US-11-172-145-6	Sequence 6, Appli
44	86	4.6	2053	7	US-11-013-759-9	Sequence 9, Appli
45	85.5	4.5	443	7	US-11-196-475-166	Sequence 166, App

ALIGNMENTS

RESULT 1

US-10-793-626-1780
; Sequence 1780, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIORITY FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1780
; LENGTH: 1155
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-1780

Query Match	7.0%;	Score 132.5;	DB 6;	Length 1155;
Best Local Similarity	22.3%;	Pred. No. 0.042;		
Matches	96;	Conservative 58;	Mismatches 169;	Indels 107; Gaps 21;
Qy	18	SRPNYNPLLVYKAKIVNK-DAANYKILNLTNNYKQIETLEKONGQDQFGREKT	76	
Db	113	SKPKANEAVVTNESKPTTEAPTNEESIAETPKTTQDSTETKNNPSLKDNLNLSST	172	
Qy	77	LLKLTAMNFFNSESLSKADLFYK-----DKPEELKKYLVKVRHTFVVLINTEGDS	130	
Db	173	TSKES-----KTDESHSTQAOQMTNKSNDLTNDSPSTQSEK-----TSSQANNDSNDQ	220	
Qy	131	D-----DGLKIYKDDYD-----KFKTPSIFVFSTKEQE-----IKE	162	
Db	221	SAPSKQLDSKPEQKVYTKFNDEPTQDVEHTTTKLTPSISTDSSVNDKQDYTRSAVAS	280	
Qy	163	LPFKDNGTEKERNIAVYNNKDNHLKFTS-OYLHOASIFHAV-----NP-YCM-PLAA	213	
Db	281	LGVDNSET-----AITNAVRNLDLKAASREQINEAIIAEALKKDFSNPDYGVDTPLAL	335	
Qy	214	TPLVDDTVIGKLRTAKINFYSLNTEGL-----DGVPAFKEGVLD-----AGGAI-----	258	
Db	336	NTSQSKNSPHKASPRMNLMSLAEPNSGKNVNDKVIKTNPTLSLKNHNHANNVWPTS	395	
Qy	259	DEQFTY--HYIKNEAI-----IELIRWNNRQNSKLSALQISGAR	298	

PRIOR FILING DATE: 2004-02-06
PRIOR APPLICATION NUMBER: 60/613,854
PRIOR FILING DATE: 2004-09-28
NUMBER OF SEQ ID NOS: 82
SOFTWARE: Patent in Ver. 3.3
SEQ ID NO 42
LENGTH: 2367
TYPE: PRT
ORGANISM: Clostridium difficile
US-11-051-453-42

Query Match 5.4%; Score 102.5; DB 7; Length 2367;
Best Local Similarity 20.6%; Pred. No. 15;
Matches 89; Conservative 65; Mismatches 153; Indels 125; Gaps 20;

QY 2 QDTSVSLDSRIQASR-PNYNPL-----LVYKTIKIV-----NKDAA 41
DB 603 PYDSI---LFORNIEDSEVAIYNPTDSEIQEIDKIRIPDRISRPKILTFIGHGKARF 659
QY 42 NYKIL-NLTVNNYKQIET---LEKDNQGDQFQKERTLLKTAMSNFNSSESLKSD 97
DB 660 NTDIFAGLDVSLSEIETAGLAKEDISPK---SIEINLLGCNMFSYSVNVVEETYPGKL 716
QY 98 LFIYKQPEELKXYLK-----VHRHTFVVLINTEG-----DNSDDGLKIYKDDYDKFKTP 147
DB 717 LLRVKDKVSELMPSMSQDSIIIVSANQYEVRIINSEGRRELLDHSGEWIN-----KBE 767
QY 148 SIFFVFSTKE-----QEIKEPKDKNGTKERNIAVYNNKDNHLKFIQYHLQASIF 201
DB 768 SIIRKDISKEYISFNPKNKIIVSKNLPSELSTLLQEIIRNNSNSDIE-----815
QY 202 HAVNPGMPLAATPLVDVTGKLTAKINIFYSLINTEGLDGVPAFKEGVDLAGGAIDEQ 261
DB 816 -----LEEKVM---LAEINVISNIETQVVE---ERIEAKSLTSDSI---853
QY 262 FTYHYKNE-AIIE-----LIRIMKNNRONSLSALQSGARDNAYTSAIECLIKRFV- 314
DB 854 ---NYIKNEFKLIESISEALCDLKQONELEDSHFISPEDISETDGFISI-----RFIN 903
QY 315 -----DRGLIIOYKN-----LSLTGSPTPQLKLELSVNITYNFSINAVSLV 355
DB 904 KETGESIFVETEKTFSEYANHITBEISKITGTFDTVNGKLVKKNVLDTTHEVNTLNA 963
QY 356 ITQDIVDYQNS 367
DB 964 FFIQSLIEYNS 975

RESULT 5
US-11-052-554A-91
Sequence 91, Application US/11052554A
Publication No. US2005028866A1
GENERAL INFORMATION:
APPLICANT: Sachdeva, et al.
TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
FILE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
FILE REFERENCE: 30853/40359A
CURRENT APPLICATION NUMBER: US/11/052,554A
CURRENT FILING DATE: 2005-02-07
PRIOR APPLICATION NUMBER: US 60/589,227
PRIOR FILING DATE: 2004-07-20
PRIOR APPLICATION NUMBER: IN 173/DEL/2004
PRIOR FILING DATE: 2004-02-06
NUMBER OF SEQ ID NOS: 763
SOFTWARE: Patent in version 3.3
SEQ ID NO 91
LENGTH: 2902
TYPE: PRT
ORGANISM: Helicobacter pylori J99
US-11-052-554A-91

Query Match 5.4%; Score 102; DB 7; Length 2902;
Best Local Similarity 19.6%; Pred. No. 22;

Matches 81; Conservative 46; Mismatches 126; Indels 160; Gaps 17;
QY 12 DSRQASRPNYNPLLVYKTIKIVNKDAANYKILNLTVNNYKQIETLEK-----62
DB 2278 NAOQFQANNLTISNOQVLEKNASFVN---NLNIQGAFNNNATQKIEVLQNLVIASNASL 2333
QY 63 -----DNGNQDQFQKTKL-----LKTAMSNFFNSSEES-----92
DB 2334 STGIYGLEVGALNMLGAIHFNLENSQTPVNPLIQVGGIILNLTOTPFMNVSVANGTY 2393
QY 93 --LKSADLFYKDKPEELKXYLVHRHTFVVLINTEGNSDD--GLKIYKDDYDKFKTPS 148
DB 2394 TLLKSSYIIDYNINPNSLSQSLKY-----TLNINGNHIEEKGVLTLYCQ-----2440
QY 149 IFFVFSTKEQEIKELFKDKG-----NTEKERNIAVYNNKDNHLKFIQYHLQHA 198
DB 2441 -----RVLLQDKGLLSVALPNSNNAQQNLSLS-----VLHNQIKMSYGNKV 2484
QY 199 SIFHAVNPGMPLAATPLVDVTGKLTAKINIFYSLINTEGLDGVPAFKEGVDLAGG-- 256
DB 2485 MDF-----TPPTLQDYIVG-----IQQSALNQ-----IEAVGGNN 2515
QY 257 AI-----DEQFTYHYIKNEAIIELI-----RIMNKNFRON-----286
DB 2516 AIKWLSTMMETKENPLFAPYILENHSLEITLGVTKDLQNTASLISNPFNNATSLLEM 2575
QY 287 -----SKLSALQSGARDNAYTSAIECLIKRFVDRG---LIQYKNLS 326
DB 2576 ASYQOTSRLTKLSDFRAREGESNFSERLELKKRFRSDPNPSEVFKYSQLS 2628

RESULT 6
US-11-051-453-41
Sequence 41, Application US/11051453
Publication No. US20050287150A1
GENERAL INFORMATION:
APPLICANT: AMBROSINO, DONNA
APPLICANT: BABCOCK, GREGORY J.
APPLICANT: BROERING, THERESA
APPLICANT: GRAZIANO, ROBERT
APPLICANT: HERNANDEZ, HECTOR JAVIER
APPLICANT: LOWY, ISRAEL
APPLICANT: MANDELL, ROBERT
APPLICANT: MOLRINE, DEBORAH
APPLICANT: THOMAS, JR., WILLIAM D.
APPLICANT: ZHANG, HUI-PEN
TITLE OF INVENTION: ANTIBODIES AGAINST CLOSTRIDIUM DIFFICILE TOXINS AND
FILE OF INVENTION: USES THEREOF
FILE REFERENCE: MJ1-001
CURRENT APPLICATION NUMBER: US/11/051,453
CURRENT FILING DATE: 2005-02-04
PRIOR APPLICATION NUMBER: 60/542,357
PRIOR FILING DATE: 2004-02-06
PRIOR APPLICATION NUMBER: 60/613,854
PRIOR FILING DATE: 2004-09-28
NUMBER OF SEQ ID NOS: 82
SOFTWARE: Patent in Ver. 3.3
SEQ ID NO 41
LENGTH: 2710
TYPE: PRT
ORGANISM: Clostridium difficile
US-11-051-453-41

Query Match 5.3%; Score 100.5; DB 7; Length 2710;
Best Local Similarity 19.7%; Pred. No. 25;
Matches 94; Conservative 63; Mismatches 180; Indels 141; Gaps 19;

QY 4 DTISVSLDSRIQAS-----RPNYINPLLVYKTIKIVNKDAANYKILNLTV 50
DB 682 DTIKLIDSPKNVENVLLGCNMFSDYFNVEETYPGKLLLSIMDKITSTLPDVNKNISITIGA 741
QY 51 NNYEKQIETLEKDNQGDQ-----FGKEKILLKTAMSN-----FFNSSESLSK-----94

Db 742 NOYEVRI-----NSEGRKELLAHSGKWINKBEAIMSDLSSEYIFFDSIDNKLKAKSKN 795
Qy 95 -----SADL-----PIYKDKPEELKYLKVRH- 116
Db 796 IFGLASISEDITKLLLDASVSPDKFILNNKLNIESSIGDYIYKLEPVKNI--IHNS 853
Qy 117 -----HTFVVLINTGD-----NSDGLKIYKDDYK-FKTPSIFFVSTKQEI-- 160
Db 854 IDDLDEFNLENVSDLEYELKLANLDEKYLISPEDISKNNSTVSVRFINKSNGESVVV 913
Qy 161 ---KELF-----KKGNTKERNIAVYNNKDNHLKFIQSYLHQASIFHAVNPGM 209
Db 914 ETEKIFSKYSHITKEISTIKNSITDVGNNLLNIQLDHTSQ-----VNTLNA 963
Qy 210 PLAATPLVD---DVTIGKLRTA-KINFVSLNLTGQVPAFKGVDLAGGAIDEQFTY 264
Db 964 AFFIOSLDIYSSNKVDLNDLSTSVKQVLAQFSTGLNTIYDSIQLVNLSNAV----- 1017
Qy 265 HYIKNEAIELIRIWNKNNRQNSKLSALQSGARDNAVYTAIECLLKRFVDRGLIIQYKN 324
Db 1018 ---NDTINVLPITTEGPIVSTILDGINL-CAAIKELLDEHDPILKKLEAKVGVLAIN 1072
Qy 325 LSLTISPTPOLKLELSVNTY-----NFSINAVSLVIT-----TQDIVDYQNSLS 369
Db 1073 MSLSTAATVASIVGIGAEVITFLLPIAGISAGIPSLVNNELILHDKATSVVNFNHL 1130

RESULT 7

US-11-055-822-1066
; Sequence 1066, Application US/11055822
; Publication No. US20050260707A1
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; FILE REFERENCE: BGI-121CPCN
; CURRENT APPLICATION NUMBER: US/11/055,822
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: 09/606,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142,101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148,613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187,970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931415.2
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931418.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1158
; SEQ ID NO 1066
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-055-822-1066

Query Match 5.2%; Score 98; DB 7; Length 701;
Best Local Similarity 20.6%; Pred. No. 6.6;
Matches 84; Conservative 41; Mismatches 95; Indels 189; Gaps 22;

Qy 1 MPQDTISVS-----LLDSRIQASRRPNYNPLLVYKTAKIKVNDK-AAANYKILNL 48
Db 354 MYEDTVNASNPIEGRIHTSNLCSEILQVSTPSEFNDDLT-----AEVGEDISCNLSLVN 409
Qy 49 TV-----NNYEKOIET-----LEKDN-----GNGO-----DOFGKEK 75
Db 410 AMAMDAPNFEKTIETIRGLTAVPEQTSIDSVPSIRKNGEAHAIGLQGMNLHGVYFGR- 468
Qy 76 TLLKTAMSNFNSSEESLSKADLF-----IYKDKPEELKYLKVRHRTFVV 121
Db 469 -----HMYGSEALDFTNAVFAAVLYOCLRASNOIATGERGPKNF----- 510
Qy 122 LINTEGDMSDGLKIYKDDYKFKTPSIFFVSTKQEIKELFKDKGNTKERNIAVYSN 181
Db 511 -----ENSKATGEYDFDDAND-----FAPSKDKVELFA----- 541
Qy 182 NKDNHLKFIQSYLHQASIFHAVNPGMPLAATPLVDDTVIGKLTAKINFYSLNLTGL 241
Db 542 -KSNIH-----TPVEDWAALK-----ADMHEGLFNR-NL 570
Qy 242 DGVPAPKEGVDLAGAIDEQFTYHYIKNEA-----IIEILIRIWNKNNRQNSKLSALQLSG 296
Db 571 QAVP-----PTGSIS-----YINNSTSIHPIASKIBI-----RKEGKIGRVYVPA 611
Qy 297 AR-----DNAYTSAIECLL-----KRFVDRGLIIQYKNLSLTL 329
Db 612 PHMDNDNLEYEFDAYEIGYEKIIDIYAVATKYVDG-----LSLTL 652

RESULT 8

US-10-512-109-27
; Sequence 27, Application US/10512109
; Publication No. US20050255546A1
; GENERAL INFORMATION:
; APPLICANT: KIRIN BEER KABUSHIKI KAISHA
; TITLE OF INVENTION: POLYPEPTIDE HAVING AN ACTIVITY TO SUPPORT PROLIFERATION OR SURVIV
; TITLE OF INVENTION: OF HEMATOPOIETIC STEM CELL OR HEMATOPOIETIC PROGENITOR CELL, AN
; TITLE OF INVENTION: FOR THE SAME
; FILE REFERENCE: 905WOIOP1572
; CURRENT APPLICATION NUMBER: US/10/512,109
; CURRENT FILING DATE: 2004-10-21
; PRIOR APPLICATION NUMBER: US 60/376,001
; PRIOR FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 27
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-512-109-27

Query Match 5.1%; Score 96; DB 6; Length 829;
Best Local Similarity 19.7%; Pred. No. 11;
Matches 70; Conservative 45; Mismatches 119; Indels 122; Gaps 15;

Qy 12 DSRIOASRPN-----YYNPLLVYKTAKIKVNDKAA-----YKILNLTVNNYEKOIETLEK 62
Db 204 DSRVALSTCNGLHGMFEDDTFVYMIETLELTDDEKSTGRPHIIQKTLAQQYKQMKNLST 263
Qy 63 DNGNGQDOF-----GKEKTLTKTAMSNFNSSEESLSKADLFYKDKPEELKYLK 113
Db 264 D---GSDQWPLLPELQWLRRRKRAVNSRGVP-----EEMKYLELMIVNDH-KTYKKHRS 314
Qy 114 VHRH-----TFVVLINTEGDSDGLKIYKDDYKFKTPSIFFV 152
Db 315 SHAHTNNFAKSVNVLVDSIYKEQLNTRVVLVAVETWTEKDHIDI-----TINPVQMLHD 368
Qy 153 FSTKEQEIKELFKDKGNTKERNIAVYNNKDNHLH-KFISQYLHQASIFH----- 202
Db 369 FSKYRQRIKQ-----HADAVHLISRTVTFHYKRSLSVYFGVCSRIR 409
Qy 203 --AVNPYGMPLAATPLVDDTVIGKLTAKINFYSLNLTGLDGVPAFK-----EGVDLAGG 256

Db 410 GGVNEXGLPMA-----VAQVLSQSLAQLGLQWEPSSRKPKECIESWGG 455
Qy 257 AIDEQFTVHYIKNEAIIELIRWKNRQNSKLSALQSGARDNAYTSAIECLLKR 312
Db 456 CIMEETGV-----SHSRKFSKCSILEY---RDFLQGGGACLFNR 492

RESULT 9

US-10-793-626-440
; Sequence 440, Application US/10793626
; Publication No. US2005025478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 440
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-440

Query Match 5.1%; Score 95.5; DB 6; Length 258;
Best Local Similarity 20.0%; Pred. No. 2.7; Mismatches 90; Indels 73; Gaps 11;
Matches 52; Conservative 45;
Qy 73 KEKTLTKTAMSNFFNSSESLKSADLFYKOKPEELKYLKVHRHTFVVLINTEGDSDD 132
Db 42 KEAVSLQTSQEGN-IHTKEVNGK---FIYASK-QDIEKAMQI-----KHSN 82
Qy 133 GLKIKDYDKPTPSIFVFSTKEQEKELFKQKNGTEKERNIAVNNKONLHLKFS 192
Db 83 DLK-TYMDISEK-----VPMSEKEVNHILKGILENGKSTFIKAQDKYEVNLYL- 131
Qy 193 QYLHOASIFHAVNPYGMPLAATPLVDDTVIGKLTAKINFYSLNETGLDGVPAKEGVD 252
Db 132 -----ISHA-----LVETGNGQSLSKGKEGNHHYNNF-----GIGAFDEDAV 171
Qy 253 LAGGAIDEG-----FTVHYIKNEAIIELIRWKNRQNSKLSALQLSG 296
Db 172 KTGKSFQKQKVTTEKAIMGGAWFRVHYFRKNQLSLYQMRWNPQNG-----QHQQY 224
Qy 297 ARDNAYTSAIECLLRFVDR 316
Db 225 ASDIQANNIADMEKYDK 244

RESULT 10

US-11-069-642-20
; Sequence 20, Application US/11069642
; Publication No. US20050260626A1
; GENERAL INFORMATION:
; APPLICANT: LORENS, JAMES B.
; APPLICANT: PRAY, TODD R.
; APPLICANT: KINGSLEA, TODD M.
; APPLICANT: BENNETT, MARK K.
; TITLE OF INVENTION: IN VIVO PRODUCTION OF CYCLIC PEPTIDES FOR
; FILE REFERENCE: RIGL-022CIP3
; CURRENT APPLICATION NUMBER: US/11/069,642
; CURRENT FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: 10/232,758
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: 10/422,536

; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 09/800,770
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/187,130
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 168
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Methanococcus jannaschii
US-11-069-642-20

Query Match 5.0%; Score 94; DB 7; Length 496;
Best Local Similarity 23.3%; Pred. No. 8.1;
Matches 57; Conservative 34; Mismatches 76; Indels 78; Gaps 12;
Qy 6 ISVSLDSRQASRPNYN-----PLAVYTKAKLVNKNDAANYKILNLTNNYKQIETL 60
Db 274 IIFSLPESYNALTAGYFDTGCFSLLYDKAKKH-----NLRMVLTSKRRLDVL 322
Qy 61 EKDNNGQDQFQKTKLLKTAMSNFFNSSESLKSADLFYKDKPEELKYLKVHRHTFV 120
Db 323 EK-IGVILNSIGILNTLHKSREVYSLISNKSLET-----FKK---JAKYLKIRKEAFI 373
Qy 121 VLINTEGDSDDGLKIYKDDY-----DKFKTPSIF--FVFSTKEQE-----IKELF 164
Db 374 -----NGYTKYKHEERFECDDLFPVKEVFKLTTEKGRKEILKDSKIHENWY 422
Qy 165 KOKGNT---EKERNIAVNNKONLHLKFSQ-----YHQASI 200
Db 423 KEKTNIPREKLTVLRYANSE--HKEFLEKIVNGDISFVRVKVENIPYDGYVDLSI 480
Qy 201 PHAVN 205
Db 481 KHNQN 485

RESULT 11

US-10-485-517-141
; Sequence 141, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynexus Incorporated
; APPLICANT: Foster, Simon
; APPLICANT: Mond, James
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P100629WO
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 141
; LENGTH: 1290
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-485-517-141

Query Match 5.0%; Score 94; DB 6; Length 1290;
Best Local Similarity 22.4%; Pred. No. 28;
Matches 66; Conservative 46; Mismatches 91; Indels 92; Gaps 16;
Qy 37 NKDAANYK-----ILNLT-VNVEKQIETLEKONGN---GDDQFQKKEKTLTKTAMSNFF 86
Db 63 NONVANAKTTAKNALNNLTSTINNAQKALKSQIEGATTVAGVQVSTTASELNTAMSNLQ 122
Qy 87 NS-SEESLSKADLFYKDKPEELKYLKVHRHTFVVLINTEGDSDDGLKIYKDDYDKFK 145

Db 123 NGINDEAATKAAL-----NGTONLEK-AKOHANTAI-----DGLS----- 156
 QY 146 TSPISFFVSTKEQIKELPKD-----KGNTEKERN-----IAVSNKNDLH 187
 Db 157 -----HLTNAQKALQVLQOSTTVAEQNGEKANNDVDAWMDKLRQSIADNATTKQON 211
 QY 188 LKFIQYLHQASIFHAVNPG-MPLAATPLVDVTI-----GKLRQAKINFYSLNLTGLD 242
 Db 212 YTDASQNKDAYNNAVTTAQGIIDQTSPTLPTVINQAAGQVSTK-----NALN--GNE 265
 QY 243 GYPAKEGVDLAGGIDQFTHYIKNEAIIELIRIWNKNRONSLSALQLSGA 297
 Db 266 NLEAAQKQASQSLGSLD-----NLNNAQKQVTD-QINGA 299

RESULT 12
 US-11-013-759-11
 ; Sequence 11, Application US/11013759
 ; Publication No. US20050249747A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Loosmore, Sheena M.
 ; APPLICANT: Sasaki, Ken
 ; APPLICANT: Yang, Yan Ping
 ; APPLICANT: Klein, Michel H.
 ; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
 ; FILE REFERENCE: 1038-921MIS:jb
 ; CURRENT APPLICATION NUMBER: US/11/013,759
 ; CURRENT FILING DATE: 2004-12-16
 ; PRIOR APPLICATION NUMBER: US/09/361,619
 ; PRIOR FILING DATE: 1999-07-27
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 11
 ; LENGTH: 2314
 ; TYPE: PRT
 ; ORGANISM: Moraxella catarrhalis
 US-11-013-759-11

Query Match 5.0%; Score 94; DB 7; Length 2314;
 Best Local Similarity 22.5%; Pred. No. 61;
 Matches 75; Conservative 47; Mismatches 138; Indels 74; Gaps 16;
 QY 31 TAKIKVNDAAVNYKILNTVNNYKQIETLEKXNGGODQFGKXTLLKTA-----MSNF 85
 Db 795 TAKVYDETNQSKV-TYDVNVDEKTIK-LTDGNGK-TNKIGVKITTLTTWANGKATNF 851
 QY 86 FNSSELSKADLFYKQKPEELKKYLK-VHR-----HTFVVLINTEGNSDDGLK 135
 Db 852 STTNDALVNA-----KDIKENLNTLAKBIHTTKGTADTALQTFKV--KKDQATDDETIT 904
 QY 136 IYKDDYDKFTPSIFVVFSTKQEIKELPKDKGNTKERNIAVSNKNDLHLKFIQYL 195
 Db 905 VGKDGTONKQTN-----TLKLG-----ENGLVATNKDGTVTFGINTQSG 946
 QY 196 HQASIFHAVNPGMPLAATPLVDVTIGK--LRTAKINFYSLNLTGLDGVPAKPE--G 250
 Db 947 LKAGSTTLUNKDGLSKNPASNEQIQVGADGVKPAKVKN--SSTGIDGTSRITKQDQIG 1004
 QY 251 VDLAGGAIHQFTHYIKNEAIIELIRIWNKN-NRQNSKLSALQSGARDNAYTSAIECL 309
 Db 1005 FTGANGSLDIT-KPHLTCKDKVGEVITNTGINAGKKITNIQ----- 1047
 QY 310 LKRFVDRGLIIQKNSLSLTSTPQLKLSVNI 343
 Db 1048 -----SGDITQNSDAVTGGRVYDLKTELESKI 1075

RESULT 13
 US-11-155-288-20
 ; Sequence 20, Application US/11155288
 ; Publication No. US2006008468A1
 ; GENERAL INFORMATION:

; APPLICANT: Chiang, Chih-Sheng
 ; APPLICANT: Simard, John J.L.
 ; TITLE OF INVENTION: COMBINATIONS OF TUMOR-ASSOCIATED
 ; TITLE OF INVENTION: ANTIGENS IN DIAGNOSTICS FOR VARIOUS TYPES OF CANCERS
 ; FILE REFERENCE: MANK.050A
 ; CURRENT APPLICATION NUMBER: US/11/155,288
 ; CURRENT FILING DATE: 2005-06-17
 ; PRIOR APPLICATION NUMBER: 60/580,969
 ; PRIOR FILING DATE: 2004-06-17
 ; NUMBER OF SEQ ID NOS: 40
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 20
 ; LENGTH: 976
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-11-155-288-20

Query Match 4.9%; Score 93.5; DB 7; Length 976;
 Best Local Similarity 22.9%; Pred. No. 21;
 Matches 50; Conservative 33; Mismatches 108; Indels 27; Gaps 6;
 QY 14 RIQASRPNNYPLLVYKTAKIKVNDAAVNYKILN---LTVNNYKQIETLEKXNGGODQ 70
 Db 605 RKQVENKMKYTEELQEQENKALK-KKGTABSKQLNAVYKIKVNDLLELSAKQKQFGEITDT 663
 QY 71 FGKERTLLKTAMSNFFNSSELSKADLF1-----YKDKPEELKKYLKVRHTFVV 121
 Db 664 YKQIEDKKEISENLLLEVEKAKVIADAVKQKEIDKRCQHKIAEMVALMEKHKHQYDK 723
 QY 122 LINTEGNSDGLKIYKDDYDKFTPSIFFVFST-----KQEIKELPKDKGNTK 172
 Db 724 II--EERDSGLGYSKEQESLSRASLEISLNLKABELLSVKYKQLEIEREKEKLEKREA 781
 QY 173 ERNIAVSNKNDLHLKFIQYLHQASIFHAVNPGMP 210
 Db 782 KENTATLKEKDKKTQTF1---LETPEIYWKLDKSKAVP 816

RESULT 14
 US-11-055-822-1070
 ; Sequence 1070, Application US/11055822
 ; Publication No. US20050260707A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pompejus, Markus
 ; APPLICANT: Kroger, Burkhard
 ; APPLICANT: Schroder, Hartwig
 ; APPLICANT: Zelder, Oskar
 ; APPLICANT: Haberhauer, Gregor
 ; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
 ; FILE REFERENCE: BGI-121CPCN
 ; CURRENT APPLICATION NUMBER: US/11/055,822
 ; CURRENT FILING DATE: 2005-02-11
 ; PRIOR APPLICATION NUMBER: 09/606,740
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: 60/141,031
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: 60/142,101
 ; PRIOR FILING DATE: 1999-07-02
 ; PRIOR APPLICATION NUMBER: 60/148,613
 ; PRIOR FILING DATE: 1999-08-12
 ; PRIOR APPLICATION NUMBER: 60/187,970
 ; PRIOR FILING DATE: 2000-03-09
 ; PRIOR APPLICATION NUMBER: DE 19930476.9
 ; PRIOR FILING DATE: 1999-07-01
 ; PRIOR APPLICATION NUMBER: DE 19931415.2
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931418.7
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931419.5
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931420.9
 ; PRIOR FILING DATE: 1999-07-08

· This Page Blank (uspto)

Search completed: January 24, 2006, 19:59:15
Job time : 4.46592 secs

This Page Blank (uspto)